Statistical aspects of the fractional stochastic calculus

Ciprian A. Tudor, Frederi Viens

Statistical Aspects of the Fractional Stochastic Calculus*

Ciprian A. Tudor[1] Frederi G. Viens[2]

[1]SAMOS-MATISSE, Centre d'Economie de La Sorbonne,

Université de Paris 1 Panthéon-Sorbonne ,

90, rue de Tolbiac, 75634, Paris, France.

tudor@univ-paris1.fr

[2] Dept. Statistics and Dept. Mathematics, Purdue University,

150 N. University St., West Lafayette, IN 47907-2067, USA.

viens@purdue.edu

May 23, 2006

Abstract

We apply the techniques of stochastic integration with respect to the fractional Brownian motion and the theory of regularity and supremum estimation for stochastic processes to study the maximum likelihood estimator (MLE) for the drift parameter of stochastic processes satisfying stochastic equations driven by fractional Brownian motion with any level of Hölder-regularity (any *Hurst* parameter). We prove existence and strong consistency of the MLE for linear

*AMS 2000 subject classifications. Primary: 62M09; secondary: 60G18, 60H07, 60H10.

Key words and phrases: maximum likelihood estimator, fractional Brownian motion, strong consistency, stochastic differential equation, Malliavin calculus, Hurst parameter.

Acknowledgment of support: second author partially supported by NSF grant no. 0204999.

and nonlinear equations. We also prove that a version of the MLE using only discrete observations is still a strongly consistent estimator.

1 Introduction

Stochastic calculus with respect to fractional Brownian motion (fBm) has recently known an intensive development, motivated by the wide array of applications of this family of stochastic processes. For example, fBm is used as a model in network traffic analysis; recent work and empirical studies have shown that traffic in modern packet-based high-speed networks frequently exhibits fractal behavior over a wide range of time scales; this has major implications for the statistical study of such traffic. An other example of applications is in quantitative finance and econometrics: the fractional Black-Scholes model has been recently introduced (see e.g. [19], [14]) and this motivates the statistical study of stochastic differential equations governed by fBm.

The topic of parameter estimation for stochastic differential equations driven by standard Brownian motion is of course not new. Diffusion processes are widely used for modeling continuous time phenomena; therefore statistical inference for diffusion processes has been an active research area over the last decades. When the whole trajectory of the diffusion can be observed, then the parameter estimation problem is somewhat simpler. But in practice data is typically collected at discrete times and thus of particular contemporary interest are works in which an approximate estimator, using only information gleaned from the underlying process in discrete time, is able to do as well as an estimator that uses continuously gathered information. This is in fact a rather challenging question and several methods have been employed to construct good estimators for discretely observed diffusions; amongs these methods, we refer to numerical approximation to the likelihood function (see Aït-Sahalia [1], Poulsen [34], Beskos et al. [4]), martingale estimating functions (see Bibby and Sorensen [5]), indirect statistical inference (see Gourieroux et al. [18]), or the Bayesian approach (see Elerian et al. [15]), some sharp probabilistic bounds on the convergence of estimators in [6], or [12], [33], [10] for particular situations. We mention the survey [38] for parameter estimation in discrete cases, further details in the works of [27], [20], or the book [22].

Parameter estimation questions for stochastic differential equations driven by fBm are, in contrast, in their infancy. Some of the main contributions include [25], [24], [36] or [26]. We take up these estimation questions in this article. Our purpose is to contribute further to the study of the statistical aspects of the fractional stochastic calculus, by introducing the systematic use of efficient tools from stochastic analysis, to yield results which hold in some non-linear generality. We consider the following stochastic equation

$$X_t = \theta \int_0^t b(X_s)ds + B_t^H, \quad X_0 = 0 \tag{1}$$

where B^H is a fBm with *Hurst* parameter $H \in (0,1)$ and the nonlinear function b satisfies some regularity and non-degeneracy conditions. We estimate the parameter θ on the basis of the observation of the whole trajectory of the process X. The parameter H, which is assumed to be known, characterizes the local behavior of the process, with Hölder-regularity increasing with H; if $H = 1/2$, fBm is standard Brownian motion (BM), and thus has independent increments; if $H > 1/2$, the increments of fBm are positively correlated, and the process is more regular than BM; if $H < 1/2$, the increments are negatively correlated, and the process is less regular than BM. H also characterizes the speed of decay of the correlation between distant increments. Estimating long-range dependence parameters is a difficult problem in itself, which has received various levels of attention depending on the context; the text [3] can be consulted for an overview of the question; we have found the yet unpublished work [11], available online, which appears to propose a good solution applicable directly to fBm. Herein we do not address the Hurst parameter estimation issue.

The results we prove in this paper are as follows:

- for every $H \in (0,1)$, we give concrete assumptions on the nonlinear coefficient b to ensure the existence of the maximum likelihood estimator (MLE) for the parameter θ (Proposition 1);
- for every $H \in (0,1)$ and under certain hypotheses on b which include non-linear classes, we prove the strong consistency of the MLE (Theorems 2 and 3, depending on whether $H < 1/2$ or $H > 1/2$; and Proposition 2 and Lemma 3 for the scope of non-linear applicability of these theorems); note that for $H > 1/2$ and b linear, this has also been proved in [24];
- for every $H \in (0,1)$, the bias and mean-square error for the MLE are estimated in the linear case (Proposition 3); this result was established for $H > 1/2$ in [24].

In this paper we also present a first practical implementation of the MLE studied herein, using only discrete observations of the solution X of equation (1), by replacing integrals with their Riemann sum approximations. We show that

- the discretization time-step for the Riemann sum approximations of the MLE can be fixed while still allowing for a strongly consistent estimator in large time, a result valid in the linear case and some non-linear classes (Proposition 5 and Theorem 4).

To establish all these results, we use techniques in stochastic analysis including the *Malliavin* calculus, and supremum estimations for stochastic processes. The Malliavin calculus, or the stochastic calculus of variations, was introduced by P. Malliavin in [29] and developed by D. Nualart in [31]. Its original purpose was to study the existence and the regularity of the density of solutions to stochastic differential equations. Since our hypotheses in the present paper to ensure existence and strong consistency of the MLE are given in terms of certain densities (see Condition (C)), the techniques of the Malliavin calculus appear as a natural tool.

We believe our paper is the first instance where the Malliavin calculus and supremum estimations are used to treat parameter estimation questions for fractional stochastic equations. These techniques should have applications and implications in statistics and probability reaching beyond the question of MLE for fBm. For example, apart from providing the first proof of strong consistency of the MLE for an fBm-driven differential equation with non-linear drift or with $H < 1/2$, more generally, in (Itô-) diffusion models, the strong consistency of an estimator follows if one can prove that an expression of the type $I_t := \int_0^t f^2(X_s)ds$ tends to ∞ as $t \to \infty$ almost surely. To our knowledge, a limited number of methods has been employed to deal with this kind of problem: for example, if X is Gaussian the Laplace transform can be computed explicitly to show that $\lim_{t\to\infty} I_t = \infty$ a.s.; also, if X is an ergodic diffusion, a local time argument can be used to show the above convergence. Particular situations have also been considered in [20], [21]. We believe that our stochastic analytic tools constitute a new possibility, judging by the fact that the case of $H < 1/2$ is well within the reach of our tools, in contrast with the other above-mentioned methods, as employed in particular in [24] (see however a general Bayesian-type problem discussed in [25]).

The organization of our paper is as follows. Section 2 contains preliminaries on the fBm. In Section 3 we show the existence of the MLE for the parameter θ in (8) and in Section 4 we study its asymptotic behavior. Section 5 contains some additional results in the case when the drift function is linear. In Section 6, a discretized version of the MLE is studied. The Appendix (Section 7) contains most of the technical proofs.

We gratefully acknowledge our debt to the insightful comments of the editor, associate editor, and two referees, which resulted in several important improvements on an earlier version of this paper.

2 Preliminaries on the fractional Brownian motion and fractional calculus

We consider $(B_t^H)_{t\in[0,T]}$, $B_0^H = 0$ a fractional Brownian motion with Hurst parameter $H \in (0,1)$, in a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. This is a centered Gaussian process with covariance function R given by

$$R(t,s) = \mathbf{E}\left[B_t^H B_s^H\right] = \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right) \qquad s,t \in [0,T]. \tag{2}$$

Let us denote by K the kernel of the fBm such that (see e.g. [30])

$$B_t^H = \int_0^t K(t,s)dW_s \tag{3}$$

where W is a Wiener process (standard Brownian motion) under $\mathbf{P}$. Denote by $\mathcal{E}_H$ the set of step functions on $[0,T]$ and let $\mathcal{H}$ be the canonical Hilbert space of the fBm; that is, $\mathcal{H}$ is the closure of $\mathcal{E}$ with respect to the scalar product

$$\langle 1_{[0,t]}, 1_{[0,s]}\rangle_{\mathcal{H}} = R(t,s).$$

The mapping $1_{[0,t]} \to B_t^H$ can be extended to a isometry between $\mathcal{H}$ and the Gaussian space generated by B^H and we denote by $B^H(\varphi)$ the image of $\varphi \in \mathcal{H}$ by this isometry.

We also introduce the operator K^* from $\mathcal{E}_H$ to $L^2([0,T])$ defined by

$$(K^*\varphi)(s) = K(T,s)\varphi(s) + \int_s^T (\varphi(r) - \varphi(s)) \frac{\partial K}{\partial r}(r,s)dr. \tag{4}$$

With this notation we have $(K^* 1_{[0,t]})(s) = K(t,s)$ and hence the process

$$W_t = \int_0^t (K^{*,-1} 1_{[0,t]})(s) dB_s^H \tag{5}$$

is a Wiener process (see [2]); in fact, it is *the* Wiener process referred to in formula (3), and for any non-random $\varphi \in \mathcal{H}$, we have $B^H(\varphi) = \int_0^T (K^*\varphi)(s)dW(s)$, where the latter is a standard Wiener integral with respect to W.

Lastly we recall some elements of fractional calculus. Let f be an L^1 function over the interval $[0,T]$ and $\alpha > 0$. Then

$$I_{0+}^{\alpha} f(t) = \frac{1}{\Gamma(\alpha)} \int_0^T \frac{f(s)}{(t-s)^{1-\alpha}} ds \quad \text{and} \quad D_{0+}^{\alpha} f(t) = \frac{1}{\Gamma(1-\alpha)} \frac{d}{dt} \int_0^T \frac{f(s)}{(t-s)^{\alpha}} ds$$

are the Riemann-Liouville fractional integrals and derivatives of order $\alpha \in (0,1)$. The latter admit the following Weil representation

$$D_{0+}^{\alpha} f(t) = \frac{1}{\Gamma(1-\alpha)} \left(\frac{f(t)}{t^{\alpha}} + \alpha \int_0^t \frac{f(t) - f(y)}{(t-y)^{\alpha+1}} dy \right)$$

where the convergence of the integrals at $t = y$ holds in the L^p-sense $(p > 1)$. We can formally define, for negative orders $(-\alpha < 0)$, the fractional integral operators as $I_{\pm}^{-\alpha} = D_{\pm}^{\alpha}$. If K_H is the linear operator (isomorphism) from $L^2([0,T])$ onto $I_{+}^{H+\frac{1}{2}}(L^2([0,T]))$ whose kernel is $K(t,s)$, [13] can be consulted for formulas for K_H; we provide here the formulas for the inverse operator of K_H in terms of fractional integrals

$$\left(K_H^{-1} h\right)(s) = s^{H-\frac{1}{2}} I_{0+}^{\frac{1}{2}-H} (s^{\frac{1}{2}-H} h'(s))(s), \qquad H \le \frac{1}{2} \tag{6}$$

and

$$\left(K_H^{-1}h\right)(s) = s^{H-\frac{1}{2}} D_{0+}^{H-\frac{1}{2}} (s^{\frac{1}{2}-H} h'(s))(s), \qquad H \geq \frac{1}{2}. \tag{7}$$

3 The maximum likelihood estimator for fBm-driven stochastic differential equations

We will analyze the estimation of the parameter $\theta \in \Theta \subset \mathbb{R}$ based on the observation of the solution X of the stochastic differential equation

$$X_t = \theta \int_0^t b(X_s)ds + B_t^H, \qquad X_0 = 0 \tag{8}$$

where B^H is a fBm with $H \in (0,1)$ and $b : \mathbb{R} \to \mathbb{R}$ is a measurable function. Let us recall some known results concerning equation (8):

- In [32] the authors proved the existence and uniqueness of a strong solution to equation (8) under the following assumptions on the coefficient b:
  - if $H < \frac{1}{2}$, b satisfies the linear growth condition $|b(x)| \leq C(1 + |x|)$,
  - if $H > \frac{1}{2}$, b is Hölder-continuous of order $\alpha \in (1 - \frac{1}{2H}, 1)$.
- In [7] an existence and uniqueness result for (8) is given when $H > \frac{1}{2}$ under the hypothesis $b(x) = b_1(x) + b_2(x)$, b_1 satisfying the above conditions and b_2 being a bounded nondecreasing left (or right) continuous function.

Remark 1 *The case of the Hölder-continuous drift is elementary: it is not difficult to show that the usual Picard iteration method can be used to prove the existence and uniqueness of a strong solution.*

Throughout the paper, from now on, we will typically avoid the use of explicit H-dependent constants appearing in the definitions of the operator kernels related to this calculus, since our main interest consists of asymptotic properties for estimators. In consequence, we will use the notation $C(H), c(H), c_H, \cdots$ for generic constants depending on H, which may change from line to line.

Our construction is based on the following observation (see [32]). Consider the process $\tilde{B}_t^H = B_t^H + \int_0^t u_s ds$ where the process u is adapted and with integrable paths. Then we can write

$$\tilde{B}_t^H = \int_0^t K(t,s) dZ_s \tag{9}$$

where

$$Z_t = W_t + \int_0^t K_H^{-1}\left(\int_0^{\cdot} u_r dr\right)(s) ds. \tag{10}$$

We have the following Girsanov theorem.

Theorem 1 *i) Assume that u is an adapted process with integrable paths such that*

$$t \to \int_0^t u_s ds \in I^{H+\frac{1}{2}}\left(L^2([0,T])\right) \ \ a.s.$$

ii) Assume that $\mathbf{E}(V_T) = 1$ where

$$V_T = \exp\left(-\int_0^T K_H^{-1}\left(\int_0^{\cdot} u_r dr\right)(s) dW_s - \frac{1}{2}\int_0^T \left(K_H^{-1}\left(\int_0^{\cdot} u_r dr\right)(s)\right)^2 ds\right). \tag{11}$$

Then under the probability measure $\tilde{\mathbf{P}}$ defined by $d\tilde{\mathbf{P}}/d\mathbf{P} = V_T$ it holds that the process Z defined in (10) is a Brownian motion and the process $\tilde{B}^H$ (9) is a fractional Brownian motion on $[0,T]$.

Hypothesis. We need to make, at this stage and throughout the remainder of the paper, the following assumption on the drift: b is differentiable with bounded derivative b'; thus the affine growth condition holds.

This Girsanov theorem is the basis for the following expression of the MLE.

Proposition 1 *Denote, for every $t \in [0,T]$, by*

$$Q_t = Q_t(X) = K_H^{-1}\left(\int_0^{\cdot} b(X_r)dr\right)(t). \tag{12}$$

Then $Q \in L^2([0,T])$ almost surely and the MLE is given by

$$\theta_t = -\frac{\int_0^t Q_s dW_s}{\int_0^t Q_s^2 ds}. \tag{13}$$

Before proving Proposition 1, we need the following estimates:

Lemma 1 *For every $s, t \in [0,T]$,*

$$\sup_{s\le t}|X_s| \le \left(Ct + \sup_{s\le t}\left|B_s^H\right|\right)e^{Kt} \tag{14}$$

and

$$|X_t - X_s| \le C(H,T,\theta)\left(1 + \sup_{u\le T}|X_u|\right)|t-s| + \left|B_t^H - B_s^H\right|. \tag{15}$$

Proof: We have, for any s

$$|X_s| \le \int_0^s |\theta|\,|b(X_u)|\,du + \left|B_s^H\right| \le C\int_0^s (1+|X_u|)\,du + \sup_{u\le s}\left|B_u^H\right|$$

and by Gronwall's lemma

$$|X_s| \le \left(Cs + \sup_{u\le s}\left|B_u^H\right|\right)e^{Cs}, \qquad s\in[0,T]$$

and the estimate (14) follows. The second estimate follows by b's affine growth. ■

Proof of Proposition 1. Let

$$h(t) = \int_0^t b(X_s)\,ds.$$

We prove that the process h satisfies i) and ii) of Theorem 1. Note first that the application of the operator K_H^{-1} preserves the adaptability. We treat separately the cases when H is bigger or less than one half.

The case $H < 1/2$. To prove i), we only need to show that $Q \in L^2([0,T])$ $\mathbf{P}$-a.s. Indeed i) is equivalent to the following, almost-surely:

$$h \in I_+^{H+1/2}\left(L^2([0,T])\right) \Longleftrightarrow K_H^{-1}h \in K_H^{-1}\left(I_+^{H+1/2}\left(L^2([0,T])\right)\right).$$

Then using the isomorphism property of K_H we see that i) is equivalent to $K_H^{-1}h \in L^2([0,T])$, which means $Q \in L^2([0,T])$ a.s. by definition. Now using relation (6) we thus have, for some constant C_H which may change from line to line, using the hypothesis $|b(x)| \leq C(1+|x|)$, for all $s \leq T$,

$$|Q_s| \leq C_H s^{H-1/2}\left|\int_0^s (s-u)^{-1/2-H} u^{1/2-H} b(X_u)\,du\right| \leq C_H\left(1+\sup_{u\leq s}|X_u|\right), \quad (16)$$

which we can rewrite, thanks to Lemma 1, as

$$\sup_{s\leq T}|Q_s| \leq C(H,T)\left(1+\sup_{s\leq T}|X(s)|\right),$$

which, thanks to inequality (14), is of course much stronger than $Q \in L^2([0,T])$ a.s., since $\sup_{s\leq T}|X(s)|$ has moments of all orders (see [32]).

To prove ii) it suffices to show that there exists a constant $\alpha > 0$ such that

$$\sup_{s\leq T}\mathbf{E}\left(\exp(\alpha Q_s^2)\right) < \infty.$$

Indeed, one can invoke an argument used by Friedman in [17], Theorem 1.1, page 152, showing that this condition implies the so-called Novikov condition (see [35]), itself implying ii). Since Q satisfies (16), the above exponential moment is a trivial consequence of inequality (14) and the Fernique's theorem on the exponential integrability of the square of a seminorm of a Gaussian process.

The case $H > 1/2$. Using formula (7) we have in this case that

$$Q_s = c_H \left[s^{\frac{1}{2}-H} b(X_s) + \left(H - \frac{1}{2}\right) s^{H-\frac{1}{2}} \int_0^s \frac{b(X_s) s^{\frac{1}{2}-H} - b(X_u) u^{\frac{1}{2}-H}}{(s-u)^{H+\frac{1}{2}}} du \right] \tag{17}$$

$$= c_H \left[s^{\frac{1}{2}-H} b(X_s) + \left(H - \frac{1}{2}\right) s^{H-\frac{1}{2}} b(X_s) \int_0^s \frac{s^{\frac{1}{2}-H} - u^{\frac{1}{2}-H}}{(s-u)^{H+\frac{1}{2}}} du \right.$$

$$\left. + \left(H - \frac{1}{2}\right) s^{H-\frac{1}{2}} \int_0^s \frac{b(X_s) - b(X_u)}{(s-u)^{H+\frac{1}{2}}} u^{\frac{1}{2}-H} du \right]$$

and using the fact that

$$\int_0^s \left(s^{\frac{1}{2}-H} - u^{\frac{1}{2}-H} \right) (s-u)^{-H-\frac{1}{2}} du = c(H) s^{1-2H}$$

we get

$$|Q_s| \le c_H \left(s^{\frac{1}{2}-H} |b(X_s)| + s^{H-\frac{1}{2}} \int_0^s \frac{b(X_s) - b(X_u)}{(s-u)^{H+\frac{1}{2}}} u^{\frac{1}{2}-H} du \right) := A(s) + B(s).$$

The first term $A(s)$ above can be treated as in [32], proof of Theorem 3, due to our Lipschitz assumption on b. We obtain that for every $\lambda > 1$,

$$\mathbf{E} \left(\exp(\lambda \int_0^t A_s^2 ds) \right) < \infty. \tag{18}$$

To obtain the same conclusion for the second summand $B(s)$ we note that by Lemma 1, up to a multiplicative constant, the random variable

$$G = \sup_{0 \le u < s \le T} \frac{|X_s - X_u|}{|u-s|^{H-\varepsilon}}$$

is bounded by

$$\left(1 + \sup_{u \le T} |X_u| \right) |t-s|^{1-H+\varepsilon} + \sup_{0 \le u < s \le T} \frac{|B_s^H - B_u^H|}{|u-s|^{H-\varepsilon}}$$

and it suffices to use the calculations contained in [32].

Conclusion. Properties i) and ii) are established for both cases of H, and we may apply Theorem 1. Expression (13) for the MLE follows a standard calculation, since (using the notation P_θ for the probability measure induced by $(X_s)_{0\le s\le t}$, and the fact that $P_0 = \mathbf{P}$),

$$F(\theta) := \log \frac{dP_\theta}{dP_0} = -\theta \int_0^t Q_s dW_s - \frac{\theta^2}{2} \int_0^t Q_s^2 ds. \tag{19}$$

■

We finish this section with some remarks that will relate our construction to previous works ([25], [24], [36]). Details about these links are given in Section 5.

Alternative form of the MLE. By (8) we can write, by integrating the quantity $K^{*,-1} 1_{[0,t]}(s)$ for s between 0 and t,

$$\int_0^t \left(K^{*,-1} 1_{[0,t]}(\cdot)\right)(s) dX_s = \theta \int_0^t \left(K^{*,-1} 1_{[0,t]}(\cdot)\right)(s) b(X_s) ds + W_t. \tag{20}$$

On the other hand, by (8) again,

$$X_t = \int_0^t K(t,s) dZ_s \tag{21}$$

where Z is given by (10). Therefore, we have the equality

$$\int_0^t \left(K^{*,-1} 1_{[0,t]}(\cdot)\right)(s) dX_s = Z_t. \tag{22}$$

By combining (20) and (22) we obtain

$$\int_0^t K_H^{-1}\left(\int_0^\cdot b(X_r) dr\right)(s) ds = \int_0^t \left(K^{*,-1} 1_{[0,t]}(\cdot)\right)(s) b(X_s) ds$$

and thus the function $t \to \int_0^t \left(K^{*,-1} 1_{[0,t]}(\cdot)\right)(s) b(X_s) ds$ is absolutely continuous with respect to the Lebesgue measure and

$$Q_t = \frac{d}{dt} \int_0^t \left(K^{*,-1} 1_{[0,t]}(\cdot)\right)(s) b(X_s) ds. \tag{23}$$

By (10) we get that the function (19) can be written as

$$F(\theta) = -\theta \int_0^t Q_s dZ_s + \frac{\theta^2}{2} \int_0^t Q_s^2 ds.$$

As a consequence, the maximum likelihood estimator θ_t has the equivalent form

$$\theta_t = \frac{\int_0^t Q_s dZ_s}{\int_0^t Q_s^2 ds}. \tag{24}$$

The above formula (24) shows explicitly that the estimator θ_t is observable if we observe the whole trajectory of the solution X.

4 Asymptotic behavior of the maximum likelihood estimator

This section is devoted to studying the strong consistency of the MLE (13). A similar result has been proven in the case $b(x) \equiv x$ and $H > \frac{1}{2}$ in [24]. We propose here a proof of strong consistency for a class of functions b which contains significant non-linear examples. By replacing (10) in (24), we obtain that

$$\theta_t - \theta = \frac{\int_0^t Q_s dW_s}{\int_0^t Q_s^2 ds}$$

with Q given by (12) or (23). To prove that $\theta_t \to \theta$ almost surely as $t \to \infty$ (which means by definition that the estimator θ_t is strongly consistent), by the strong law of large numbers we need only show that

$$\lim_{t\to\infty} \int_0^t Q_s^2 ds = \infty \text{ a.s. } . \tag{25}$$

To prove that $\lim_{t\to\infty} \int_0^t Q_s^2 ds = \infty$ in a non-linear case, it is necessary to make some assumption of non-degeneracy on the behavior of b. In order to illustrate our method using the least amount of technicalities, we will restrict our study to the case where the function $|b|$ satisfies a simple probabilistic estimate with respect to fractional Brownian motion.

(C) There exist positive constants t_0 and K_b, both depending only on H and the function b, and a constant $\gamma < 1/(1+H)$ such that for all $t \geq t_0$ and all $\varepsilon > 0$, we have $\tilde{\mathbf{P}}\left[\left|Q_t(\tilde{\omega})\right|/\sqrt{t} < \varepsilon\right] \leq \varepsilon t^{\gamma H} K_b$, where under $\tilde{\mathbf{P}}$, $\tilde{\omega}$ has the law of fractional Brownian motion with parameter H.

4.1 The case $H < \frac{1}{2}$

In this part we prove the following result.

Theorem 2 *Assume that $H < 1/2$ and that Condition (C) holds. Then the estimator θ_t is strongly consistent, that is,*

$$\lim_{t\to\infty} \theta_t = \theta \textit{ almost surely.}$$

Before proving this theorem, we discuss Condition (C). To understand this condition, we first note that with μ_H^t the positive measure on $[0,t]$ defined by $\mu_H^t(dr) = (r/t)^{1/2-H}(t-r)^{-1/2-H}\,dr$, according to the representation (6), we have

$$Q_t = \int_0^t \mu_H^t(ds)\, b(\tilde{\omega}_s)$$

and therefore, by the change of variables $r = s/t$,

$$\frac{Q_t}{\sqrt{t}} = \int_0^1 \mu_H^1(dr)\,\frac{b(\tilde{\omega}_{tr})}{t^H} \tag{26}$$

$$\stackrel{\mathcal{D}}{=} \int_0^1 \mu_H^1(dr)\,\frac{b\left(t^H\tilde{\omega}_r\right)}{t^H}, \tag{27}$$

where the last inequality is in distribution under $\tilde{\mathbf{P}}$.

If b has somewhat of a linear behavior, we can easily imagine that $b\left(t^H\tilde{\omega}_r\right)/t^H$ will be of the same order as $b(\tilde{\omega}_r)$. Therefore $Q_t/\sqrt{t}$ should behave, in distribution for fixed t, similarly to the universal random variable $\int_0^1 \mu_H^1(dr)\, b(\tilde{\omega}_r)$ (whose distribution depends only on b and H). Generally speaking, if this random variable has a bounded density, the strongest version of condition (C), i.e. with $\gamma = 0$, will follow. In the linear case, of course, the factors t^H disappear from expression (27), leaving a random variable which is indeed known to have has a bounded density, uniformly in t, by the arcsine law. The presence of the factor $t^{\gamma H}$ in Condition (C) gives even more flexibility, however, since in particular it allows a bound on the density of $Q_t/\sqrt{t}$ to be proportional to t^H.

Leaving aside these vague considerations, we now give, in Proposition 2, a simple sufficient condition on b which implies condition (C). The proof of this condition uses the tools of the Malliavin calculus; as such, it requires some extra regularity on b. We also give a class of non-linear examples of b's satisfying (C) (Condition (29) in Lemma 3) which are more restricted in their global behavior than in Proposition 2, but do not require any sort of local regularity for b.

Proposition 2 *Assume $H < 1/2$. Assume that b' is bounded and that b'' satisfies $|b''(x)| \leq b_1/\left(1 + |x|^{\beta}\right)$ for some $\beta \in (H/(1-H), 1)$. Assume moreover that $|b'|$ is bounded below by a positive constant b_0. Then, letting $\gamma = 1 - \beta$, Condition (C) holds.*

Remark 2 *The condition $\gamma < 1/(1+H)$ from Condition (C) does translates as $\beta > H/(1-H)$, which is consistent with $\beta < 1$ because $H < 1/2$.*

The non-degeneracy condition on $|b'|$ above can be relaxed. It is possible, for example, to prove that if, for $x \geq x_0$, only the condition $|b'(x)| \geq x^{-\alpha}$ holds, then Condition (C) holds as long as α does not exceed a positive constant $\alpha_0(H)$ depending only on H. However, such a proof is more technical than the one given below, and we omit it.

The hypothesis of fractional power decay on b'', while crucial, does allow b to have a truly non-linear behavior. Compare with Lemma 3 below, which would correspond *to the case $\beta = 1$ here.*

The hypotheses of the above proposition imply that b is monotone.

The proof of Proposition 2 requires a criterion from the Malliavin calculus, which we present here. The book [31] by D. Nualart is an excellent source for proofs of the results we quote. Here we will only need to use the following properties of the Malliavin derivative D with respect to W (recall that W is the standard Brownian motion used in the representation (3), i.e. defined in (5)). For simplicity of notation we assume that all times are bounded by $T = 1$. The operator D, from a subset of $L^2(\Omega)$ into $L^2(\Omega \times [0,1])$, is essentially the only one which is consistent with the following two rules:

1. Consider a centered Gaussian random variable $Z \in L^2(\Omega)$; it can be therefore represented as $Z = W(f) = \int_0^1 f(s)\, dW(s)$ for some non-random function $f \in L^2([0,1])$. The operator D picks out the function f, in the sense that for any $r \in [0,1]$,
$$D_r Z = f(r).$$
2. D is compatible with the chain rule, in the sense that for any $\Phi \in C^1(\mathbf{R})$ such that both $F := \Phi(Z)$ and $\Phi'(Z)$ belongs to $L^2(\Omega)$, for any $r \in [0,1]$,
$$D_r F = D_r \Phi(Z) = \Phi'(Z)\, D_r Z = \Phi'(Z)\, f(r).$$

For instance, using these two rules, definition (3) and formula (5) relative to the fBm $\tilde{\omega}$ under $\tilde{\mathbf{P}}$, we have that under $\tilde{\mathbf{P}}$, for any $r \leq s$,
$$D_r b\left(t^H \tilde{\omega}_s\right) = t^H b'\left(t^H \tilde{\omega}_s\right) K(s,r). \tag{28}$$

It is convenient to define the domain of D as the subset $\mathbf{D}^{1,2}$ of r.v.'s $F \in L^2(\Omega)$ such that $D.F \in L^2(\Omega \times [0,1])$. Denote the norm in $L^2([0,1])$ by $\|\cdot\|$. The set $\mathbf{D}^{1,2}$

forms a Hilbert space under the norm defined by

$$\|F\|_{1,2}^2 = \mathbf{E}\,|F|^2 + \mathbf{E}\,\|D.F\|^2 = \mathbf{E}\,|F|^2 + \mathbf{E}\int_0^1 |D_r F|^2\,dr.$$

Similarly, we can define the second Malliavin derivative D^2F as a member of $L^2\left(\Omega \times [0,1]^2\right)$, using an iteration of two Malliavin derivatives, and its associated Hilbert space $\mathbf{D}^{2,2}$. Non-Hilbert spaces, using other powers than 2, can also be defined. For instance, the space $\mathbf{D}^{2,4}$ is that of random variables F having two Malliavin derivatives, and satisfying

$$\begin{aligned}\|F\|_{2,4}^4 &= \mathbf{E}\,|F|^4 + \mathbf{E}\,\|D.F\|^2 + \mathbf{E}\left\|D^2_{\cdot,\cdot}F\right\|^4_{L^2([0,1]^2)} \\ &= \mathbf{E}\,|F|^4 + \mathbf{E}\int_0^1 |D_r F|^2\,dr + \mathbf{E}\left(\int_0^1 |D_r D_s F|^2\,drds\right)^2 < \infty\end{aligned}$$

We also note that the so-called *Ornstein-Uhlenbeck* operator L acts as follows (see [31, Proposition 1.4.4]):

$$LF = L\Phi(Z) = -Z\Phi'(Z) + \Phi''(Z)\,\|f\|^2\,.$$

We have the following Lemma, whose proof we omit because it follows from ([31, Proposition 2.1.1. and Exercise 2.1.1]).

Lemma 2 *Let F be a random variable in $\mathbf{D}^{2,4}$, such that $\mathbf{E}\left[\|DF\|^{-8}\right] < \infty$. Then F has a continuous and bounded density f given by*

$$f(x) = \mathbf{E}\left[1_{(F>x)}\left(\frac{-LF}{\|DF\|^2} - 2\frac{\langle DF \otimes DF; D^2F\rangle_{L^2([0,1]^2)}}{\|DF\|^4}\right)\right]$$

Proof of Proposition 2.

Step 0: strategy. Using the identity in law (27), and the shorthand notation $\mu = \mu_1^H$, let

$$F = \frac{Q_t}{\sqrt{t}} = \int_0^1 \mu(dr)\,\frac{b\left(t^H\tilde{\omega}_r\right)}{t^H}.$$

It is sufficient to prove that F has a density which is bounded by $K_b t^{\gamma H}$ where the constant K_b depends only on b and H. Indeed $\tilde{\mathbf{P}}\left[\left|Q_t\left(\tilde{\omega}\right)\right|/\sqrt{t}<\varepsilon\right] \leq \int_0^{\varepsilon} K_b t^{\gamma H} dx = \varepsilon t^{\gamma H} K_b$. In this proof, $C_{b,H}$ denotes a constant depending only on b and H, whose value may change from line to line.

Step 1: calculating the terms in Lemma 2. We begin with the calculation of DF. Since the Malliavin derivative is linear, we get $D_r F = t^{-H}\int_0^1 \mu\left(ds\right) D_r\left(b\left(t^H\tilde{\omega}_s\right)\right)$. Then from (28) we get

$$D_r F = \int_r^1 \mu\left(ds\right) b'\left(t^H\tilde{\omega}_s\right) K\left(s,r\right).$$

Thus we can calculate

$$\begin{aligned}\left\|DF\right\|^2 &= \int_0^1 dr \left|\int_r^1 \mu\left(ds\right) b'\left(t^H\tilde{\omega}_s\right) K\left(s,r\right)\right|^2 \\ &= \int_0^1\int_0^1 \mu\left(ds\right)\mu\left(ds'\right) b'\left(t^H\tilde{\omega}_s\right) b'\left(t^H\tilde{\omega}_{s'}\right)\int_0^{\min(s,s')} K\left(s,r\right)K\left(s,r'\right)dr \\ &= \int_0^1\int_0^1 \mu\left(ds\right)\mu\left(ds'\right) b'\left(t^H\tilde{\omega}_s\right) b'\left(t^H\tilde{\omega}_{s'}\right) R\left(s,s'\right),\end{aligned}$$

where R is the covariance of fBm in (2). A similar calculation yields

$$D^2_{q,r}F = t^H\int_{\max(q,r)}^1 \mu\left(ds\right) b''\left(t^H\tilde{\omega}_s\right) K\left(s,r\right)K\left(s,q\right)$$

and

$$\left\|D^2F\right\|^2_{L^2([0,1]^2)} = t^{2H}\int_0^1\int_0^1 \mu\left(ds\right)\mu\left(ds'\right) b''\left(t^H\tilde{\omega}_s\right) b''\left(t^H\tilde{\omega}_{s'}\right)\left|R\left(s,s'\right)\right|^2.$$

For the Ornstein-Uhlenbeck operator, which is also linear, we get

$$-LF = \int_0^1 \mu\left(ds\right)\left(b'\left(t^H\tilde{\omega}_s\right)\tilde{\omega}_s + b''\left(t^H\tilde{\omega}_s\right)t^H s^{2H}\right).$$

Step 2: estimating the terms in Lemma 2. With the expressions in the previous step, using the hypotheses of the proposition, we now obtain, for some constant C_H

depending only on H,

$$\tilde{\mathbf{E}}\left[|LF|\right] \leq C_H \left(\left\|b'\right\|_\infty + t^H b_1 \int_0^1 \mu\left(ds\right) s^{2H} \tilde{\mathbf{E}}\left[\frac{1}{1+t^{\beta H}\left|\tilde{\omega}_s\right|^\beta}\right]\right)$$
$$= C_H \left(\left\|b'\right\|_\infty + t^H b_1 \mathbf{E}\left[\int_0^1 \mu\left(ds\right) s^{2H} \frac{1}{1+\left(ts\right)^{\beta H}\left|Z\right|^\beta}\right]\right),$$

where Z is a generic standard normal random variable. We deal first with the integral in s. Using the fact that $\mu\left(ds\right)$ has a bounded density, and the elementary fact that for any $a>0$ and any $\alpha<1$, we have $\int_0^1 ds\left(1+as^\alpha\right)^{-1} \leq \left(1-\alpha\right)^{-1} a^{-1}$, we may now write, using $a = t^{\beta H}\left|Z\right|^\beta$ and $\alpha = \beta H$,

$$\tilde{\mathbf{E}}\left[|LF|\right] \leq C_{H,b}\left(1+t^{H(1-\beta)}\frac{1}{1-\beta H}\mathbf{E}\left[\left|Z\right|^{-\beta}\right]\right)$$
$$\leq C_{H,b}\left(1+t^{H(1-\beta)}\right).$$

The estimation of $\left\|D^2 F\right\|$ is similar. Using its expression in the previous step, the boundedness of $d\mu/ds \cdot d\mu/ds' \cdot R\left(s,s'\right)$, and the fact that $\int_0^1 ds\left(1+as^\alpha\right)^{-2} \leq \left(1-2\alpha\right)^{-1} a^{-2}$, with $\alpha = \beta H < 1/2$, we get

$$\tilde{\mathbf{E}}\left\|D^2 F\right\|_{L^2([0,1]^2)} \leq C_{b,H} t^{H(1-\beta)}.$$

Also almost surely, for any $p \geq 2$, for some constant $C_{H,p}$ depending only on H and p, since b' has a constant sign, we obtain

$$\frac{1}{\left\|DF\right\|^p} = \left(\iint_{[0,1]^2} \mu\left(ds\right)\mu\left(ds'\right) R\left(s,s'\right)\left|b'\left(t^H\tilde{\omega}_s\right) b'\left(t^H \tilde{\omega}_{s'}\right)\right|\right)^{-p/2} \leq C_{H,p} b_0^{-p}.$$

Lastly, it is convenient to invoke the Cauchy-Schwartz inequality to get

$$\frac{\left\langle DF\otimes DF; D^2F\right\rangle_{L^2([0,1]^2)}}{\left\|DF\right\|^4} \leq \frac{\left\|D^2F\right\|_{L^2([0,1]^2)}}{\left\|DF\right\|^2},$$

Step 3: applying Lemma 2; conclusion. The third estimate in the previous step (for $p=8$) proves trivially that $\tilde{\mathbf{E}}\left\|DF\right\|^{-8}$ is finite. That $F \in \mathbf{D}^{2,4}$ follows again trivially from the boundedness of b' and b'' using the expressions in Step 1. Thus Lemma 2 applies. We conclude from the estimates in the previous step that F has a density f which is bounded as

$$f(x) \le C_{H,b}\left(1 + t^{H(1-\beta)}\right) b_0^{-2}$$

With $t \ge 1$, the conclusion of the proposition follows. ■

A smaller class of functions b satisfying condition (C), but which is not restricted to $H < 1/2$, is given in the following result, proved in the Appendix.

Lemma 3 *Let $H \in (0,1)$. Assume $xb(x)$ has a constant sign for all $x \in \mathbf{R}_+$ and a constant sign for all $x \in \mathbf{R}_-$. Assume*

$$|b(x)/x| = c + h(x) \tag{29}$$

for all x, where c is a fixed positive constant, and $\lim_{x\to\infty} h(x) = 0$. Then Condition (C) is satisfied with $\gamma = 0$.

Condition (C) also holds for any b of the above form to which a constant C is added: $|(b(x) - C)/x| = c + h(x)$ and $\lim_{x\to\infty} h(x) = 0$. Note that this condition is less restrictive than saying b is asymptotically affine, since it covers the family $b(x) = C + cx + (|x| \wedge 1)^{\alpha}$ for any $\alpha \in (0,1)$. In some sense, Condition (C) with $\gamma = 0$ appears to be morally equivalent to this class of functions.

Proof of Theorem 2. Since we only want to show that (25) holds, and since $\int_0^t |Q_s|^2\, ds$ is increasing, it is sufficient to satisfy condition (25) for t tending to infinity along a sequence $(t_n)_{n\in\mathbf{N}}$. We write, according to the representation (6), for each fixed $t \ge 0$,

$$I_t = I_t(X) := \int_0^t |Q_s(X)|^2\, ds = \int_0^t \left| \int_0^s \mu_H^s(dr)\, b(X_r) \right|^2 ds$$

where X is the solution of the Langevin equation (8) and the positive measure μ_H^s is defined by $\mu_H^s(dr) = (r/s)^{1/2-H}(s-r)^{-1/2-H}\,dr$. Recall from the Girsanov Theorem 1 applied to X, that with

$$\eta_T = \exp\left(-\int_0^T Q_s(X)\,dW_s - \frac{1}{2}\int_0^T |Q_s(X)|^2\,ds\right)$$

where W is standard Brownian motion under $\mathbf{P}$, we have that under the probability measure $\tilde{\mathbf{P}}$ defined by its Radon-Nikodym derivative $\frac{d\tilde{\mathbf{P}}}{d\mathbf{P}}\Big|_{\mathcal{F}_T^X} = \eta_T$ for all $T \geq 0$, X is a fractional Brownian motion with parameter H. Moreover, since η_T is a true martingale, this Girsanov transformation can be reversed. See [35, Theorem VIII.1.7]: with $L_t = \int_0^T Q_s(X)\,dW_s$, we can write that

$$\frac{d\mathbf{P}}{d\tilde{\mathbf{P}}}\bigg|_{\mathcal{F}_T^X} = \tilde{\eta}_T,$$

where under $\tilde{\mathbf{P}}$, $\tilde{\eta}$ can be written as the exponential martingale

$$\tilde{\eta}_t = \exp\left(\tilde{L}_t - \frac{1}{2}\left\langle \tilde{L}\right\rangle_t\right)$$

for some martingale $\tilde{L}$ under $\tilde{\mathbf{P}}$ satisfying $\left\langle \tilde{L}\right\rangle_t = \langle L\rangle_t = I_t(X) = \int_0^t |Q_s(X)|^2\,ds$. Here, since X is still a fractional Brownian motion with parameter H under $\tilde{\mathbf{P}}$, we will use the notation $\tilde{\omega}$ for X, to signify that X does not have the law of X under $\mathbf{P}$.

Thence consider a sequence of constants $(\beta_n)_{n=0}^\infty$ which will be chosen later. We claim that for any such β_n and for $t = t_n = n^k$ with fixed $k \geq 1$, it holds that

$$\mathbf{P}\left[I_{t_n}(X) < \beta_{t_n}\right] \leq C(p,H,b)\frac{1}{n^{2/p}}\exp\left(\lambda\beta_{t_n}\right) \tag{30}$$

with some $p > 1$, p close to 1, where the constant $C(p,H,b)$ depends only on p, H, and b. The proof of (30) is very technical and will be given in the Appendix, Section 7.2. It uses the fact, proved therein, that the supremum $\sup_{s\in[0,t]}|Q_s|$ behaves like a sub-Gaussian random variable with scale $\sqrt{t}$, meaning that after dividing by $\sqrt{t}$, the tail of its distribution is no heavier than a standard normal's; it is also based on the following crucial lemma, whose proof is likewise postponed to Section 7.2; this lemma will be used in other parts of this paper as well. The notation $b_n \ll t_n$ means $\lim_{n\to\infty} b_n/t_n = 0$.

Lemma 4 *Let $V_t := t^{-1/2}Q_t$. The process V is $\tilde{\mathbf{P}}$-almost surely continuous. Moreover, if $b_n > 0$ and $b_n \ll t_n$ for large n, then for any $M > 2$, there exists a constant $C_{M,H}$ such that*

$$\tilde{\mathbf{E}}\left[\sup_{s,t\in[t_n-b_n,t_n]} |V_t - V_s|^M\right] \leq C_{M,H,b}\left(\frac{b_n}{t_n}\right)^{HM}.$$

In order for the bound (30) to be summable in n, it is sufficient to choose $\beta_{t_n} = (2\lambda)^{-1}\log n$, and to take p very close to 1, so that $n^{-2/p}\exp(\lambda\beta_{t_n})$ can be bounded above by any power $n^{-3/2+\varepsilon}$ for any $\varepsilon > 0$. Therefore, by the Borel-Cantelli lemma, there exists a random almost surely finite integer $n_0(\tilde{\omega})$ depending on the function $b(\cdot)$ via the constant K_b, and depending on the constants H, p, and λ, such that for all $n > n_0(\tilde{\omega})$,

$$I_{n^k}(\tilde{\omega}) \geq \frac{1}{2\lambda}\log n,$$

where the constants p and λ can be chosen, for instance, as described in the lines following inequality (50), implying the result (25) along the sequence $t_n = n^k \to \infty$, and the Theorem. ■

4.2 The case $H > \frac{1}{2}$

Due to the fact that the function Q is less regular in this case, we should not expect that the proof of the following Theorem be a consequence of the proof in the case $H < 1/2$. Nevertheless, it deviates from the former proof very little. On the other hand, we cannot rely on Proposition 2 to find a convenient sufficient condition for Condition (C); instead one can look to the non-linear class of examples in Lemma 3, which satisfy the strong version ($\gamma = 0$) of Condition (C) for all $H \in (0, 1)$. The next result's proof is in the Appendix.

Theorem 3 *Assume that $H > \frac{1}{2}$ and b satisfies Condition (C) with $\gamma = 0$ (e.g. b satisfies Condition (29) in Lemma 3). Then the maximum likelihood estimator θ_t is strongly consistent.*

5 The linear case

In this section we present some comments in the case when the drift b is linear. We will assume that $b(x) \equiv x$ to simplify the presentation. In this case, the solution X to equation (8) is the fractional Ornstein-Uhlenbeck process and it is possible to prove more precise results concerning the asymptotic behavior of the maximum likelihood estimator.

Remark 3 *In [9], it is shown that there exists an unique almost surely continuous process X that satisfies the Langevin equation (8) for any $H \in (0, 1)$. Moreover the process X can be represented as*

$$X_t = \int_0^t e^{\theta(t-u)} dB_u^H, \quad t \in [0, T] \tag{31}$$

where the above integral is a Wiener integral with respect with B^H (which exists also as a pathwise Riemann-Stieltjes integral). It follows from the stationarity of the increments of B^H that X is stationary and the decay of its auto-covariance function is like a power function. The process X is ergodic, and for $H > \frac{1}{2}$, it exhibits a long-range dependence.

Let us briefly recall the method employed in [24] to estimate the drift parameter of the fractional OU process. Let us consider the function , for $0 < s < t \leq 1$,

$$k(t,s) = c_H^{-1} s^{\frac{1}{2}-H}(t-s)^{\frac{1}{2}-H} \text{ with } c_H = 2H\Gamma(\frac{3}{2}-H)\Gamma(H+\frac{1}{2}) \tag{32}$$

and let us denote its Wiener integral with respect to B^H by

$$M_t^H = \int_0^t k(t,s)dB_s^H. \tag{33}$$

It has been proved in [30] that M^H is a Gaussian martingale with bracket

$$\langle M^H \rangle_t := \omega_t^H = \lambda_H^{-1} t^{2-2H} \text{ with } \lambda_H = \frac{2H\Gamma(3-2H)\Gamma(H+\frac{1}{2})}{\Gamma(\frac{3}{2}-H)}. \tag{34}$$

The authors called M^H the fundamental martingale associated to fBm. Therefore, observing the process X given by (8) is the same thing as observing the process

$$Z_t^{KB} = \int_0^t k(t,s)dX_s$$

which is actually a semimartingale with the decomposition

$$Z_t^{KB} = \theta \int_0^t Q_s^{KB} d\omega_s^H + M_t^H \tag{35}$$

where

$$Q_t^{KB} = \frac{d}{d\omega^H}\int_0^t k(t,s)X_s ds, \qquad t \in [0,T]. \tag{36}$$

By using Girsanov's theorem (see [30] and [24]) we obtain that the MLE is given by

$$\theta_t := \theta_t^{KB} = \frac{\int_0^t Q_s^{KB} dZ_s^{KB}}{\int_0^t (Q_s^{KB})^2 d\omega_s^H}. \tag{37}$$

Remark 4 *We can observe that our operator (13) or (24) coincides (possibly up to a multiplicative constant) with the one used in [24] and given by (37). Assume that* $H < \frac{1}{2}$*; the case* $H > \frac{1}{2}$ *is just a little more technical.*

Proof. Using relations (12) and (32) we can write

$$Q_t = C(H)t^{H-\frac{1}{2}} \int_0^t s^{\frac{1}{2}-H}(t-s)^{-\frac{1}{2}-H} b(X_s)ds$$
$$= C(H)t^{H-\frac{1}{2}} \int_0^t \frac{d}{dt} k(t,s)b(X_s)ds = C(H)t^{H-\frac{1}{2}} \frac{d}{dt}\int_0^t k(t,s)b(X_s)ds.$$

It is not difficult to see that $\frac{d}{dt}\int_0^t k(t,s)b(X_s)ds = C(H)t^{1-2H}Q_t^{KB}$ and therefore

$$Q_t = C(H)t^{1/2-H}Q_t^{KB}. \tag{38}$$

On the other hand, it can be similarly seen that

$$Z_t^{KB} = C(H)\int_0^t s^{\frac{1}{2}-H} dZ_s. \tag{39}$$

and the estimation given by (37) and (24) coincide up to a constant. ■

To compute the expression of the bias and of the mean square error and to prove the strong consistency of the estimator, one has the option, in this explicit linear situation, to compute the Laplace transform of the quantity $\int_0^t (Q_s^{KB})^2 d\omega_s^H$. This is done for $H > 1/2$ in [24], Section 3.2, and the following properties are obtained:

- the estimator θ_t is strongly consistent, that is,

$$\theta_t \to \theta \text{ almost surely when } t \to \infty,$$

- the *bias* and the *mean square error* are given by
  - If $\theta < 0$, when $t \to \infty$, then

$$\mathbf{E}(\theta_t - \theta) \backsim \frac{2}{t}\ ,\ \mathbf{E}(\theta_t-\theta)^2 \backsim \frac{2}{t}|\theta|, \tag{40}$$

  - If $\theta > 0$, when $t \to \infty$, then

$$\mathbf{E}(\theta_t - \theta) \backsim -2\sqrt{\pi \sin \pi H}\theta^{\frac{3}{2}} e^{-\theta t}\sqrt{t} \tag{41}$$

$$\mathbf{E}(\theta_t - \theta)^2 \backsim 2\sqrt{\pi \sin \pi H}\theta^{\frac{5}{2}} e^{-\theta t}\sqrt{t}. \tag{42}$$

It is interesting to realize that the rate of convergence of the bias and of the mean square error does not depends on H. In fact the only difference between the classical case (see [28]) and the fractional case is the presence of the constant $\sqrt{\pi H}$ in (40), (41), and (42). It is natural to expect the same results if $H < \frac{1}{2}$. This is true, as stated below, and proved in the Appendix in Section 7.4.

Proposition 3 *If $H < \frac{1}{2}$, then (40), (41), and (42) hold.*

6 Discretization

In this last section we present a discretization result which allows the implemention of an MLE for an fBm-driven stochastic differential equation.

We first provide background information on discretely observed diffusion processes in the classical situation when the driving noise is the standard Brownian motion. Assume that

$$dX_t = b(X_t, \theta) + \sigma(X_t, \theta)dW_t$$

where σ, b are known functions, W is a standard Wiener process and θ is the unknown parameter. If continuous information is available, the parameter estimation by using maximum likelihood method is somewhat simpler; indeed, the maximum likelihood function can be obtained by means of the standard Girsanov theorem and there are results on the asymptotic behavior of the estimator (consistency, efficiency etc...). We refer to the monographs [8], [37] or [23] for complete expositions of this topic.

"Real-world" data is, however, typically discretely sampled (for example stock prices collected once a day, or, at best, at every tick). Therefore, statistical inference for discretely observed diffusion is of great interest for practical purposes and at the same time, it poses a challenging problem. Here the main obstacle is the fact that discrete-time transition functions are not known analytically and consequently the likelihood function is in general not tractable. In this situation there are alternative methods to treat the problem. Among these methods, we refer to numerical approximation to the likelihood function (see Aït-Sahalia [1], Poulsen [34], Beskos et al. [4]), martingale estimating functions (see Bibby and Sorensen [5]), indirect statistical inference (see Gourieroux et al. [18]) or Bayesian approaches (see Elerian et al. [15]). We refer to [38] for a survey of methods of estimations in the discrete case. When the transition functions of the diffusion X are known, and $\sigma(x,\theta) = \sigma x$ with σ unknown and not depending on θ, then Dacunha-Castelle and Florens-Zmirou [12] propose a maximum likelihood estimator which is strongly consistent for the pair (θ, σ). They also gives a measure of the loss of information due to the discretization as a function depending on the interval between two observations.

A more particular situation is the case when σ is known (assume that $\sigma = 1$). Then the maximum likelihood function, given by $\exp\left(\theta \int_0^t b(X_s)dX_s - \frac{\theta^2}{2}\int_0^t b(X_s)^2 ds\right)$, can been approximated using Riemann sums as

$$\exp\left(\theta \sum_{i=0}^{N-1} b(X_{t_i})\left(X_{t_{i+1}} - X_{t_i}\right) - \frac{\theta^2}{2}\sum_{i=0}^{N-1} b(X_{t_i})^2 (t_{i+1} - t_i)\right).$$

As a consequence a the following maximum likelihood estimator can be obtained from the discrete observations of the process X at times $t_0, \ldots, t_N$ in a fixed interval $[0,T]$,with discrete mesh size decreasing to 0 as $N \to \infty$ (see [27], including proof of convergence to the continuous MLE):

$$\theta_{N,T} = -\frac{\sum_{i=0}^{N-1} b(X_{t_i})(X_{t_{i+1}} - X_{t_i})}{\sum_{i=0}^{N-1} \left|b(X_{t_i})\right|^2 (t_{i+1} - t_i)}. \tag{43}$$

In the fractional case, we are aware of no such results. We propose a first concrete solution to the problem. We choose to work with the formula (24) by replacing the stochastic integral in the numerator and the Riemann integral in the denominator by their corresponding approximate Riemann sums, using discrete integer time. Specifically we define for any integer $n \geq 1$,

$$\bar{\theta}_n := \frac{\sum_{m=0}^{n} Q_m \left(Z_{m+1} - Z_m\right)}{\sum_{m=0}^{n} \left|Q_m\right|^2}. \tag{44}$$

Our goal in this section is to prove that $\bar{\theta}_n$ is in fact a consistent estimator for θ. By our Theorems 2 and 3, it is of course sufficient to prove that $\lim_{n\to\infty} \left(\bar{\theta}_n - \theta_n\right) = 0$ almost surely. One could also consider the question of the discretization of θ_T using a fine time mesh for fixed T, and showing that this discretization converges almost surely to θ_t; by time-scaling such a goal is actually equivalent to our own.

It is crucial to note that in the fractional case the process Q given by (12) depends continuously on X and therefore the discrete observation of X does not allow directly to obtain the discrete observation of Q. We explain how to remedy this issue: Q_m appearing in (44) can be easily approximated if we know the values of $X_n, n \geq 1$ since only a deterministic integral appears in the expression of (12); indeed, for $H < \frac{1}{2}$, the quantity

$$\check{Q}_n = c(H) n^{H-\frac{1}{2}} \sum_{j=0}^{n-1} (n-j)^{-H-\frac{1}{2}} j^{\frac{1}{2}-H} b(X_j) \tag{45}$$

can be deduced from observations and it holds that $\lim_n (Q_n - \check{Q}_n) = 0$ almost surely. This last fact requires a proof, which is simpler than the proof of convergence of $\bar{\theta}_n - \theta_n$ to 0, but still warrants care; we present the crucial estimates of this proof in the appendix, in Section 7.5.1.

Note moreover that calculation of $\bar{\theta}_n$ also relies on Z_m, which is not observable; yet from formula (22), where Z_m is expressed as a stochastic integral of a deterministic function against the increments of X, again, we may replace all the Z_m's by their Riemann sum; proving that these sums converge to the Z_m's follows from calculations which are easier than those presented in Section 7.5.1, because they only require discretizing the deterministic integrand. We summarize this discussion in the following precise statement, referring to Section 7.5.1 for indications of its proof.

Proposition 4 *With $\check{Q}_n$ as in (45) and $\check{Z}_n = \sum_{j=0}^{n-1} \left(K^{*,-1}1_{[0,n]}(\cdot)\right)(j)\,(X_{j+1} - X_j)$, then almost surely $\bar{\theta}_n - \check{\theta}_n$ converges to 0, where $\check{\theta}_n$ is given by (44) with Z and Q replaced by $\check{Z}$ and $\check{Q}$.*

Let $\langle M \rangle_n$ denote the quadratic variation at time n of a square-integrable martingale M. We introduce the following two semimartingales:

$$A_t := \int_0^t Q_s dZ_s \tag{46}$$

$$B_t := \int_0^t Q_{[s]} dZ_s \tag{47}$$

where $[s]$ denotes the integer part of s. We clearly have $B_n = \sum_{m=0}^{n-1} Q_m\,(Z_{m+1} - Z_m)$. Thus using the fact that Z is a Brownian motion under $\tilde{\mathbf{P}}$, we see that

$$\langle B \rangle_n = \sum_{m=0}^{n-1} |Q_m|^2 \tag{48}$$

while

$$\langle A - B \rangle_n = \int_0^n \left|Q_s - Q_{[s]}\right|^2 ds = \sum_{m=0}^{n-1} \int_m^{m+1} |Q_s - Q_m|^2\, ds. \tag{49}$$

Therefore from definitions (13) and (44) we immediately get the expressions

$$\theta_n = \frac{A_n}{\langle A \rangle_n} \text{ and } \bar{\theta}_n = \frac{B_n}{\langle B \rangle_n}.$$

The following proposition defines a strategy for proving that $\bar{\theta}_n$ – and, by the previous proposition, $\check{\theta}_n$ – is a consistent estimator for θ. See the Appendix, Section 7.5.2, for its proof.

Proposition 5 *Let $H \in (0,1)$. If there exists a constant $\alpha > 0$ such that*

- *$n^{\alpha} \langle A - B \rangle_n / \langle B \rangle_n$ is bounded almost surely for n large enough,*
- *for all $k \geq 1$, for some constant $K > 0$, almost surely, for large n, $\langle B \rangle_n^k \geq K\mathbf{E}\left[\langle B \rangle_n^k\right]$,*
- *and for all $k > 1$, $\mathbf{E}\left[\left|\langle A - B \rangle_n\right|^k\right] \leq n^{-k\alpha}\mathbf{E}\left[\left|\langle B \rangle_n\right|^k\right]$,*

then almost surely $\lim_{n\to\infty} \bar{\theta}_n = \theta$.

This proposition allows us to prove the following, under the condition (C') below, which is stronger than (C), but still allows for non-linear examples.

Theorem 4 *Assume b' is bounded and the following condition holds:*

(C') *There exist positive constants t_0 and K_b, both depending only on H and the function b, such that for all $t \geq t_0$ and all $\varepsilon > 0$, we have $\tilde{\mathbf{P}}\left[\left|Q_t(\tilde{\omega})\right| / \sqrt{t} < \varepsilon\right] \leq \varepsilon K_b$, where under $\tilde{\mathbf{P}}$, $\tilde{\omega}$ has the law of fractional Brownian motion with parameter H.*

Then for all $H \in (0, 1/2)$, almost surely $\lim_{n\to\infty} \bar{\theta}_n = \theta$ where the discretization $\bar{\theta}_n$ of the maximum likelihood estimator θ_n is defined in (44). If $H \in (1/2, 1)$, the same conclusion holds if we assume in addition that b'' is bounded.

By Proposition 4, the above statements hold with $\bar{\theta}$ replaced by $\check{\theta}$.

Remark 5 *Condition (C') holds as soon as the random variable $Q_t(\tilde{\omega}) / \sqrt{t}$ has a density that is bounded uniformly t. When $H < 1/2$, this is a statement about the random variables $\int_0^1 \mu_1^H(ds)\, b\left(t^H \tilde{\omega}_s\right) t^{-H}$. In all cases, Condition (C') holds for the class of non-linear functions defined in Lemma 3.*

Remark 6 *We conjecture that Theorem (4) holds if we replace (C') by (C), in view for example of the fact that the conditions of Proposition 5 hold for any $\alpha < 2H$. Step 1 in the theorem's proof is the obstacle to us establishing this.*

Proof of Theorem 4: strategy and outline.

First note that since the probability measures $\mathbf{P}$ and $\tilde{\mathbf{P}}$ are equivalent (see Theorem 1), almost sure statements under one measure are equivalent to statements about the same stochastic processes under the other measure, and therefore we may prove the statements in the theorem by assuming that the process Z in the definitions (46) and (47) is a standard Brownian motion, since such is its law under $\tilde{\mathbf{P}}$. Furthermore, for the same reason, we can assume that, in these same definitions, Q is given by formula (12) where X is replaced by $\tilde{\omega}$ whose law is that of standard fBm. We will use specifically, instead of (12), the explicit formula (26) when $H < 1/2$. For $H > 1/2$ the formula (17) must be used instead, which shows the need for a control of b's second derivative. For the sake of conciseness, we restrict our proofs to the case $H < 1/2$. The result of the theorem is establised as soon as one can verify the hypotheses of Proposition 5. Here we present only the proof of the first of the three hypotheses. The other two are proved using similar or simpler techniques. To achieve our goal in this proof, it is thus sufficient to prove that almost surely, for large n, $\langle B\rangle_n \geq n^{\alpha_1}$ while $\langle A-B\rangle_n \leq n^{\alpha_2}$ where the values α_1 and α_2 are non random and $\alpha_1 > \alpha_2$. We establish these estimates in the appendix. Summarizing, we first prove that for appropriately chosen constants A and γ, for all integers m' in the interval $[n^A - n^{4\gamma A}; n^A]$, almost surely for large n, $|Q_{m'}|$ exceeds $n^{A(1/2-\gamma)}/2$ (this is done by combining Condition (C') and Lemma 4, in Step 1 in Section 7.5.3). Because of the positivity of all the terms in the series (48), this easily implies the lower bound $\langle B\rangle_n \geq cn^{2+2\gamma}$ for some constant c (Step 2 in Section 7.5.3). Step 3 in Section 7.5.3 then details how to use Lemma 4 to show the generic term of the series (49) defining $\langle A-B\rangle_n$ is bounded above by $n^{1-2\delta}$. The theorem then easily follows (Step 4 in Section 7.5.3). ■

7 Appendix

7.1 Proof of Lemma 3

Define

$$V_t := t^{-H} \int_0^1 \mu_H^1 (dr)\, b\left(t^H \tilde{\omega}_r\right) \stackrel{\mathcal{D}}{=} \frac{Q_t}{\sqrt{t}},$$

Our assumption implies four different scenarios in terms of the constant sign of b on $\mathbf{R}_+$ or $\mathbf{R}_-$. We will limit this proof to the situation where $b(x)$ has the same sign as x. The other three cases are either similar or easier. Thus we have $b(x) = cx + xh(x)$. Also define $V_* = \int_0^1 \mu_H^1 (dr)\, c\tilde{\omega}_r$ and $E_t = V_t - V_*$ so that $V_t = E_t + V_*$. Now

$$\begin{aligned} E_t &= \int_0^1 \mu_H^1 (dr) \left(b\left(t^H \tilde{\omega}_r\right) t^{-H} - c\tilde{\omega}_r\right) \\ &= t^{-H} \int_0^1 \mu_H^1 (dr) \left(b\left(t^H \tilde{\omega}_r\right) - ct^H \tilde{\omega}_r\right) \\ &= \int_0^1 \mu_H^1 (dr)\, \tilde{\omega}_r h\left(t^H \tilde{\omega}_r\right). \end{aligned}$$

For $\tilde{\mathbf{P}}$-almost every $\tilde{\omega}$, the function $\tilde{\omega}$ is continuous, and thus bounded on $[0,1]$. Therefore, $\tilde{\mathbf{P}}$-almost surely, uniformly for every $r \in [0,1]$, $\lim_{t\to\infty} h\left(t^H \tilde{\omega}_r\right) = 0$. Thus the limit is preserved after integration against μ_H^1, which means that $\tilde{\mathbf{P}}$-almost surely, $\lim_{t\to\infty} E_t = 0$.

Now fix $\varepsilon > 0$. There exists $t_0(\tilde{\omega})$ finite $\tilde{\mathbf{P}}$-almost surely such that for any $t > t_0(\tilde{\omega})$, $|E_t| \le \varepsilon$. Thus if $|V_t| < \varepsilon$, we must have $|V_*| = |V_* + E_t - E_t| = |V_t - E_t| \le |V_t| + |E_t| \le 2\varepsilon$. This proves that $\tilde{\mathbf{P}}$-almost surely,

$$\limsup_{t\to\infty} \{|V_t| < \varepsilon\} \le \{|V_*| < 2\varepsilon\}$$

and therefore,

$$\limsup_{t\to\infty} \tilde{\mathbf{P}}\left[|V_t| < \varepsilon\right] \le \tilde{\mathbf{P}}\left[\limsup_{t\to\infty} \{|V_t| < \varepsilon\}\right] \le \tilde{\mathbf{P}}\left[|V_*| < 2\varepsilon\right].$$

Now we invoke the fact that V_* is precisely the random variable studied in the first, linear, example, so that $\tilde{\mathbf{P}}\left[|V_*| < 2\varepsilon\right] \leq 2K\varepsilon$ for some constant K depending only on c and H, finishing the proof of the lemma under Condition (29).

The proof of the last statement of the lemma is identical to the above development, the constant C only adding a term to E_t which converges to 0 deterministically.■

7.2 Proof of relation (30)

Using the trivial fact that $\mathbf{1}_{(-\infty,a]}(x) \leq \exp(-\lambda x)\exp(\lambda a)$, and Hölder's inequality, we can write

$$\begin{aligned}
\mathbf{P}\left[I_t(X) < \beta_t\right] &= \tilde{\mathbf{E}}\left[\mathbf{1}_{I_t(\tilde{\omega})<\beta_t}\tilde{\eta}_t(\tilde{\omega})\right] \\
&\leq \exp(\lambda\beta_t)\,\tilde{\mathbf{E}}\left[\exp\left(-\lambda I_t(\tilde{\omega}) + \tilde{L}_t - 2^{-1}I_t(\tilde{\omega})\right)\right] \\
&= \exp(\lambda\beta_t)\,\tilde{\mathbf{E}}\left[\exp\left(-(\lambda-\nu)\,I_t(\tilde{\omega}) + \tilde{L}_t - \left(\nu+2^{-1}\right)I_t(\tilde{\omega})\right)\right] \\
&\leq \exp(\lambda\beta_t)\,\tilde{\mathbf{E}}\left[\exp\left(-p(\lambda-\nu)\,I_t(\tilde{\omega})\right)\right]^{1/p} \\
&\quad\times \tilde{\mathbf{E}}\left[\exp\left(q\tilde{L}_t - q\left(\nu+2^{-1}\right)I_t(\tilde{\omega})\right)\right]^{1/q}. \qquad (50)
\end{aligned}$$

where $0 < \nu < \lambda$ are arbitrary fixed positive constants. We may now choose the conjugate Hölder exponents $p^{-1}+q^{-1} = 1$. It will be convenient to allow $p > 1$ to be as close to 1 as possible, hence q will be very large. We also want $q^2/2 = q\left(\nu+2^{-1}\right)$. This forces us to take $\nu = 2^{-1}(q-1)$, which will also be very large. We then take λ to be a fixed value $> \nu$. The choice on q means that the last term in (50) above is equal to 1. Hence, letting

$$y := p(\lambda-\nu)$$

we have

$$\mathbf{P}\left[I_t(X) < \beta_t\right] \leq \exp(\lambda\beta_t)\,\tilde{\mathbf{E}}\left[\exp\left(-yI_t(\tilde{\omega})\right)\right]^{1/p}. \qquad (51)$$

To evaluate the above expectation, since $\exp\left(-yI_t\left(\tilde{\omega}\right)\right)$ is a random variable in the interval $(0,1)$, we first write

$$\tilde{\mathbf{E}}\left[\exp\left(-yI_t\left(\tilde{\omega}\right)\right)\right]=\int_0^1\tilde{\mathbf{P}}\left[\exp\left(-yI_t\right)>x\right]dx=\int_0^\infty e^{-z}\tilde{\mathbf{P}}\left[I_t<\frac{z}{y}\right]dz.$$

Now let $t=t_n=n^k$ for some fixed $k\geq 1$, and for all $n\in\mathbf{N}$. We also introduce a positive sequence b_n whose definition will be motivated below. We write

$$\begin{aligned}
I_{t_n}&=\int_0^{t_n}\left|Q_s\left(\tilde{\omega}\right)\right|^2 ds\\
&\geq\int_{t_n-b_n}^{t_n}\left|Q_s\left(\tilde{\omega}\right)\right|^2 ds\\
&\geq b_n\left|Q_{t_n}\left(\tilde{\omega}\right)\right|^2-\int_{t_n-b_n}^{t_n}\left|Q_{t_n}\left(\tilde{\omega}\right)-Q_s\left(\tilde{\omega}\right)\right|\left|Q_{t_n}\left(\tilde{\omega}\right)+Q_s\left(\tilde{\omega}\right)\right|ds\\
&\geq b_n\left(\left|Q_{t_n}\left(\tilde{\omega}\right)\right|^2-\sup_{s\in[t_n-b_n,t_n]}\left|Q_{t_n}\left(\tilde{\omega}\right)-Q_s\left(\tilde{\omega}\right)\right|\left|Q_{t_n}\left(\tilde{\omega}\right)+Q_s\left(\tilde{\omega}\right)\right|\right).\\
&\geq b_n\left(\left|Q_{t_n}\left(\tilde{\omega}\right)\right|^2-2\sup_{s\in[0,t_n]}\left|Q_s\left(\tilde{\omega}\right)\right|\sup_{s\in[t_n-b_n,t_n]}\left|Q_{t_n}\left(\tilde{\omega}\right)-Q_s\left(\tilde{\omega}\right)\right|\right).
\end{aligned}\tag{52}$$

We will need the result of the next lemma in order to control the variations of Q on the interval $[t_n-b_n,t_n]$. It can be considered as a consequence of the fact that $V_t:=t^{-1/2}Q_t$ is an asymptotically sub-stationary process in the second Gaussian chaos, although the proof we present below only requires the use of moments of V via the Kolmogorov continuity lemma, because of the fact that we are working in the Hölder scale of fractional Brownian regularity. Recall the statement of Lemma 4.

[Lemma 4] *Let $V_t\left(\tilde{\omega}\right):=t^{-1/2}Q_t\left(\tilde{\omega}\right)$. If $b_n>0$ and $b_n\ll t_n$, then for any $M>2$, there exists a constant $C_{M,H}$ such that $\tilde{E}\left[\sup_{s,t\in[t_n-b_n,t_n]}\left|V_t-V_s\right|^m\right]\leq C_{M,H,b}\left(\frac{b_n}{t_n}\right)^{HM}$.*

The proof of this lemma will be given further below. We now use it as follows.

Let $x = z/(2y)$. Let

$$Z_n = 2 \sup_{s\in[0,t_n]} |Q_s(\tilde{\omega})| \sup_{s\in[t_n-b_n,t_n]} |Q_{t_n}(\tilde{\omega}) - Q_s(\tilde{\omega})| / t_n.$$

We also introduce another positive sequence a_n. From (52), we have

$$\begin{aligned}
&\tilde{\mathbf{P}}\left[I_{t_n} < \frac{z}{y}\right] \\
&\leq \tilde{\mathbf{P}}\left[|Q_{t_n}(\tilde{\omega})|^2 / t_n - Z_n < \frac{x}{t_n b_n}\right] \\
&= \tilde{\mathbf{P}}\left[|Q_{t_n}(\tilde{\omega})|^2 / t_n - Z_n < \frac{x}{b_n t_n}; Z_n \geq a_n\right] \\
&+ \tilde{\mathbf{P}}\left[|Q_{t_n}(\tilde{\omega})|^2 / t_n - Z_n < \frac{x}{b_n t_n}; Z_n < a_n\right] \\
&\leq \tilde{\mathbf{P}}[Z_n \geq a_n] + \tilde{\mathbf{P}}\left[|Q_{t_n}(\tilde{\omega})|^2 / t_n < \frac{x}{b_n t_n} + a_n\right]. \qquad (53)
\end{aligned}$$

By condition (C), the last term in line (53) above can be bounded as

$$\tilde{\mathbf{P}}\left[|Q_{t_n}(\tilde{\omega})|^2 / t_n < \frac{x}{b_n t_n} + a_n\right] \leq K_b t_n^{\gamma H} \sqrt{\frac{x}{b_n t_n} + a_n}. \qquad (54)$$

We now show that the first term in line (54) is bounded as follows, for any value $M \geq 1$, for some constant $C_{M,H,b}$ depending only on M, H, and the function b:

$$\tilde{\mathbf{P}}[Z_n \geq a_n] \leq C''_{M,H,b}\left(\left(\frac{b_n}{t_n}\right)^H \frac{1}{a_n}\right)^M. \qquad (55)$$

We start off by writing,

$$\begin{aligned}
Z_n &= 2 \sup_{s\in[0,t_n]} |Q_s(\tilde{\omega})| \sup_{s\in[t_n-b_n,t_n]} |Q_{t_n}(\tilde{\omega}) - Q_s(\tilde{\omega})| / t_n. \\
&\leq 2 \sup_{s\in[0,t_n]} \frac{|Q_s(\tilde{\omega})|}{\sqrt{t_n}} \sup_{s\in[t_n-b_n,t_n]} \frac{|Q_{t_n}(\tilde{\omega}) - Q_s(\tilde{\omega})|}{\sqrt{s}} \\
&\leq 2 \sup_{s\in[0,t_n]} \frac{|Q_s(\tilde{\omega})|}{\sqrt{t_n}} \sup_{s\in[t_n-b_n,t_n]} \left(\left|\frac{Q_{t_n}(\tilde{\omega})}{\sqrt{t_n}} - \frac{Q_s(\tilde{\omega})}{\sqrt{s}}\right| + |Q_{t_n}(\tilde{\omega})| \frac{t_n - s}{\sqrt{s t_n}\left(\sqrt{s} + \sqrt{t_n}\right)}\right) \\
&\leq 2\left(1 + \frac{b_n}{2(t_n - b_n)}\right) \sup_{s\in[0,t_n]} \frac{|Q_s(\tilde{\omega})|}{\sqrt{t_n}} \sup_{s\in[t_n-b_n,t_n]} \left|\frac{Q_{t_n}(\tilde{\omega})}{\sqrt{t_n}} - \frac{Q_s(\tilde{\omega})}{\sqrt{s}}\right| \\
&\leq 4(t_n)^{-1/2} \sup_{s\in[0,t_n]} |Q_s(\tilde{\omega})| \sup_{s,t\in[t_n-b_n,t_n]} |V_t(\tilde{\omega}) - V_s(\tilde{\omega})| \qquad (56)
\end{aligned}$$

where we used the bound $s \geq t_n - b_n > t_n/2$, which holds for n large since $b_n \ll t_n$. To control the term involving $Q_s(\tilde{\omega})$, note that by (26), we get

$$\begin{aligned}
\sup_{s\in[0,t_n]} |Q_s(\tilde{\omega})| &\leq C_b \sup_{s\in[0,t_n]} \sqrt{s} \int_0^1 \mu_H^1(dr)\left(1+|\tilde{\omega}_{sr}|\right) s^{-H} dr \\
&\leq C(b,H) \sup_{s\in[0,t_n]} \left(1+\sup_{u\in[0,s]} |\tilde{\omega}_u|\right) s^{1/2-H} \\
&\leq C(b,H)\, t_n^{1/2-H} \left(1+\sup_{s\in[0,t_n]} |\tilde{\omega}_s|\right) \\
&\overset{\mathcal{D}}{=} t_n^{1/2} C(b,H) \left(t_n^{-H}+\sup_{u\in[0,1]} |\tilde{\omega}_u|\right).
\end{aligned}$$

Here the last equality is in distribution, using scaling. Now it is known that the supremum of fBm on $[0,1]$ is a subgaussian random variable with mean c_H and scale σ_H, two constants depending only on H; this fact can be proven using the standard theory of Gaussian supremum estimates (see e.g. [16]). This means that it has moments of all orders, which depend only on H.

We now apply this result and Lemma 4 to (56), to obtain

$$\begin{aligned}
&\tilde{\mathbf{P}}\left[Z_n \geq a_n\right] \\
&\leq \tilde{\mathbf{P}}\left[4\,(t_n)^{-1/2} \sup_{s\in[0,t_n]} |Q_s(\tilde{\omega})| \sup_{s,t\in[t_n-b_n,t_n]} |V_t(\tilde{\omega})-V_s(\tilde{\omega})| \geq a_n\right] \\
&\leq \tilde{\mathbf{E}}^{1/2}\left[\sup_{s,t\in[t_n-b_n,t_n]} |V_t(\tilde{\omega})-V_s(\tilde{\omega})|^{2M}\right] \tilde{\mathbf{E}}^{1/2}\left[\sup_{s\in[0,t_n]} |Q_s(\tilde{\omega})|^{2M}\right] \left(a_n\sqrt{t_n}/4\right)^{-M} \\
&\leq C'_{M,H,b} \left(\frac{b_n}{t_n}\right)^{HM} (t_n)^{M/2}\, \tilde{\mathbf{E}}^{1/2}\left[\left(1+\sup_{s\in[0,1]} |\tilde{\omega}_s|\right)^{2M}\right] \left(a_n\sqrt{t_n}/4\right)^{-M} \\
&\leq C''_{M,H,b} \left(\left(\frac{b_n}{t_n}\right)^{H} \frac{1}{a_n}\right)^{M},
\end{aligned}$$

from which our claim (55) follows. We now choose positive numbers j, k, ℓ and define

$t_n = n^k$, $b_n = n^{-j}$, $a_n = n^{-\ell}$. Thus using (53), (54), and (55),

$$\tilde{\mathbf{P}}\left[I_{t_n} < \frac{z}{y}\right] \leq K_b t_n^{\gamma H - 1/2} b_n^{-1/2} \sqrt{x + a_n b_n t_n} + C_{M,H,b}\left(\left(\frac{b_n}{t_n}\right)^H \frac{1}{a_n}\right)^M. \quad (57)$$

We are ready to show how to choose the parameters j, k, ℓ, M. First we force $a_n b_n t_n = 1$, so that $\ell = k - j$. Next we impose conditions to make summable terms in n appear on the right-hand side of (57):

- To make the second term on the right-hand side of (57) smaller than n^{-2}, since we may choose M as large as we want, we simply need to choose the parameters j and k so that the power of n in $(b_n/t_n)^H / a_n = n^{(1-H)k-(1+H)j}$ is negative, i.e.:

$$k < j\frac{1+H}{1-H}.$$

- For the first term on the right-hand side of (57), which is proportional to $t_n^{\gamma H-1/2} b_n^{-1/2} \sqrt{x+1} = n^{-k(1/2-\gamma H)+j/2}\sqrt{x+1}$, if we want to make a term n^{-2} appear, it is sufficient to require

$$k > \frac{j+4}{1-2\gamma H}.$$

To ensure that the above two restrictions on k are compatible, it is sufficient to require

$$\frac{j+4}{1-2\gamma H} < j\frac{1+H}{1-H}$$

$$\Longleftrightarrow$$

$$j\left(\frac{1+H}{1-H} - \frac{1}{1-2\gamma H}\right) > \frac{4}{1-2\gamma H}.$$

We can comply with this restriction by making j larger than $\frac{4}{1-2\gamma H}\left(\frac{1+H}{1-H} - \frac{1}{1-2\gamma H}\right)^{-1}$ as soon as the second factor in this expression is positive, which is equivalent to

$$\gamma < \frac{1}{1+H};$$

this explains why we had to include such a restriction as part of our hypothesis in Condition C.

With these choices of parameters, inequality (57) now becomes, for some other constant $C'''_{H,b}$,

$$\tilde{\mathbf{P}}\left[I_{t_n} < \frac{z}{y}\right] \leq C''_{H,b}\frac{1}{n^2}\left(1+\sqrt{x+1}\right)$$

Hence we finally obtain from (51)

$$\begin{aligned}\mathbf{P}\left[I_{t_n}(X) < \beta_{t_n}\right] &\leq \exp\left(\lambda\beta_{t_n}\right)\left[\int_0^\infty e^{-z}\tilde{\mathbf{P}}\left[I_{t_n} < 2x\right]dz\right]^{1/p} \\ &\leq \left(C''_{H,b}\right)^{1/p}\frac{1}{n^{2/p}}\exp\left(\lambda\beta_{t_n}\right)\left(1+\int_0^\infty e^{-z}\sqrt{1+\frac{z}{2p\left(\lambda-\nu\right)}}\right)^{1/p} \\ &= C\left(p,H,b\right)\frac{1}{n^{2/p}}\exp\left(\lambda\beta_{t_n}\right)\end{aligned}$$

where the constant $C(p,H,b)$ can be chosen as depending only on p, H, and b. ∎

7.2.1 Proof of Lemma 4

To apply the Kolmogorov continuity lemma (see [35, Theorem I.2.1]), we must evaluate the moments of the increments of V: let $M > 2$ and $s < t$ with $s,t \in [t_n - b_n, t_n]$; abbreviate $\mu := \mu_H^1$; let c_H denote the mass of μ, or other constants depending only on H. We have

$$\begin{aligned}&\tilde{\mathbf{E}}\left[\left|V_t - V_s\right|^M\right] \\ &= \tilde{\mathbf{E}}\left[\left|\int_0^1 \mu\left(dr\right)\left(\frac{b\left(\tilde{\omega}_{tr}\right)}{t^H} - \frac{b\left(\tilde{\omega}_{sr}\right)}{s^H}\right)\right|^M\right] \\ &\leq \left(c_H\right)^M \int_0^1 \mu\left(dr\right)\tilde{\mathbf{E}}\left[\left|\frac{b\left(\tilde{\omega}_{tr}\right)}{t^H} - \frac{b\left(\tilde{\omega}_{sr}\right)}{s^H}\right|^M\right] \\ &\leq \left(c_H\right)^M \int_0^1 \mu\left(dr\right)\left\{t^{-HM}\tilde{\mathbf{E}}\left[\left|b\left(\tilde{\omega}_{tr}\right) - b\left(\tilde{\omega}_{sr}\right)\right|^M\right] + \left(\frac{|t-s|}{s^{1+H}}\right)^M \tilde{\mathbf{E}}\left[\left|b\left(\tilde{\omega}_{sr}\right)\right|^M\right]\right\}.\end{aligned}$$

Now we use the fact that b is Lipshitz, so for some constant b_0, $|b(x) - b(y)| \leq b_0|x-y|$ and $|b(x)| \leq b_0(1+|x|)$; and we use the Gaussian law of $\tilde{\omega}$. For some constant $C_{M,H,b}$ which may change from line to line,

$$\tilde{\mathbf{E}}\left[|V_t - V_s|^M\right] \leq C_{M,H,b} \int_0^1 \mu(dr) \left\{ t^{-HM} b_0^M |t-s|^{MH} + \left(\frac{|t-s|}{s^{1+H}}\right)^M s^{MH} b_0^M \right\}$$
$$= C_{M,H,b} \int_0^1 \mu(dr) \left\{ \left|\frac{t-s}{t}\right|^{MH} + \left(\left|\frac{t-s}{s}\right|\right)^M \right\}.$$

Since we assume that $b_n \ll t_n$, we can certainly use the fact that $t_n > 2b_n$, so that $2s \geq 2(t_n - b_n) > t_n \geq t$, and therefore $(t-s)/t \leq (t-s)/s \leq 2(t-s)/t_n$ for n large enough. Hence we have proved that

$$\tilde{\mathbf{E}}\left[|V_t - V_s|^M\right] \leq C_{M,H,b} \int_0^1 \mu(dr) \left|\frac{t-s}{t_n}\right|^{MH} = C_{M,H,b} \frac{|t-s|^{MH}}{t_n^{MH}}.$$

If we now define U on the interval $[0,1]$ by $U_u = V_{t_n - b_n + ub_n}$, we see that U satisfies

$$\tilde{\mathbf{E}}\left[|U_u - U_v|^M\right] \leq C_{M,H,b} (b_n/t_n)^{MH} |u-v|^{HM}.$$

Temporarily normalizing U by the constant $C_{M,H,b}(b_n/t_n)^{MH}$, and applying [35, Theorem I.2.1], we finally get that the unnormalized U has a continuous version, and for any $\alpha < H - 1/M$ and a universal constant K

$$\tilde{\mathbf{E}}\left[\sup_{u,v\in[0,1]} \left(\frac{|U_u - U_v|}{|u-v|^\alpha}\right)^M\right] \leq K C_{M,H,b} (b_n/t_n)^{MH}.$$

Since $1 \leq |u-v|^{-\alpha}$, the statement of the Lemma follows. ■

7.3 Proof of the Theorem 3

Recall from the proof of Proposition 1 that we can write

$$Q_t = c(H) t^{\frac{1}{2}-H} b(X_t) + c'(H) \int_0^t \mu_H^t(dr) \left(b(X_t) - b(X_r)\right). \tag{58}$$

We note that in this case the expression $\mu_H^t(dr)$ does not determine a measure, but we still use this notation to simplify the presentation; the Lipschitz assumption on b and the Hölder property of X do ensure the existence of the integral.

One can actually follow the proof in the case $H < \frac{1}{2}$ line by line. All we have to do here is to prove an equivalent of Lemma 4 on the variations of Q, this being the only point where the form of Q, which differs depending on whether H is bigger or less than 1/2, is used. We will illustrate how the second summand of Q in (58) (which is the most difficult to handle) can be treated.

Denoting by $Q'_t = \int_0^t \mu_H^t(dr)\,(b(X_t) - b(X_r))$, it holds

$$
\begin{aligned}
Q'_t t^{-\frac{1}{2}} &= t^{-\frac{1}{2}} \int_0^t \mu_H^t(dr)\,(b(X_t) - b(X_r)) \\
&\stackrel{\mathcal{D}}{=} \int_0^1 \mu_H^1(dr) \frac{b(t^H \tilde{\omega}_1) - b(t^H \tilde{\omega}_r)}{t^H}.
\end{aligned}
$$

where again $\stackrel{\mathcal{D}}{=}$ denotes equality in distribution. Now, if $V'_t := t^{-\frac{1}{2}} Q'_t$, we have

$$
\begin{aligned}
V'_t - V'_s = &\int_0^1 \mu_H^1(dr) t^{-H} \left(b(t^H \tilde{\omega}_1) - b(t^H \tilde{\omega}_r) - b(s^H \tilde{\omega}_1) + b(s^H \tilde{\omega}_r) \right) \\
&+ (t^{-H} - s^{-H}) \int_0^1 \mu_H^1(dr) \left(b(s^H \tilde{\omega}_1) - b(s^H \tilde{\omega}_r) \right) := J_1 + J_2;
\end{aligned}
$$

it holds that

$$
\begin{aligned}
\tilde{\mathbf{E}} \left| J_2 \right|^M &\leq |t^{-H} - s^{-H}|^M b_0^M s^{HM} \tilde{\mathbf{E}} \left(\int_0^1 \mu_H^1(dr) \left| \tilde{\omega}_r - \tilde{\omega}_1 \right| \right)^M \\
&\leq C_{M,H,b} \left(\frac{|t-s|}{s^{1+H}} \right)^M s^{HM}
\end{aligned}
$$

and it has been already proved that this is bounded by $C_{M,H,b} \left(|t-s|/t_n \right)^{MH}$. For the term denoted by J_1 one can obtain the same bound simply by using the Lipschitz property of b: for every a, b, using the mean-value theorem, for points $\xi_a \in [s^H a, t^H a]$ and $\xi_b \in [s^H b, t^H b]$, we obtain

$$
\begin{aligned}
&\left| b(t^H a) - b(t^H b) - b(s^H a) + b(s^H b) \right| \\
&= \left| b'(\xi_a) \left(t^H - s^H \right) a + b'(\xi_b) \left(t^H - s^H \right) b \right| \\
&\leq \left\| b' \right\|_\infty |t^H - s^H| \left(|a| + |b| \right),
\end{aligned}
$$

and thus

$$\begin{aligned}
\tilde{\mathbf{E}}\,|J_1|^M &\leq C_{M,b,H}\left(\frac{|t^H - s^H|}{t^H}\right)^M \tilde{\mathbf{E}}\left(\int_0^1 \mu_H^1(dr)\left(|\tilde{\omega}_r| + |\tilde{\omega}_1|\right)\right)^M \\
&\leq C_{M,H,b}\left(\frac{|t-s|}{t_n}\right)^{MH}.
\end{aligned}$$

■

7.4 Proof of Proposition 3

To avoid tedious calculations with fractional integrals and derivatives, we will take advantage of the calculations performed in [24] when $H > \frac{1}{2}$; nevertheless we believe that a direct proof is also possible. Actually the only moment when the authors of [24] use the fact that H is bigger than $\frac{1}{2}$ is the computation of the process Q. By relations (21) and (23) we can write

$$\begin{aligned}
Q_t &= \frac{d}{dt}\int_0^t \left(K^{*,-1}1_{[0,t]}(\cdot)\right)(s)X_s ds \\
&= \frac{d}{dt}\int_0^t \left(K^{*,-1}1_{[0,t]}(\cdot)\right)(s)\left(\int_0^s K(s,v)dZ_v\right)ds \\
&= \frac{d}{dt}\int_0^t \int_v^t \left(K^{*,-1}1_{[0,t]}(\cdot)\right)(s)K(s,v)dsdZ_v
\end{aligned}$$

Note that from the formulas presented in Section 2, we have

$$\left(K^{*,-1}1_{[0,t]}(\cdot)\right)(s) = c(H)s^{\frac{1}{2}-H}\int_s^t u^{\frac{1}{2}-H}(u-s)^{-H-\frac{1}{2}}, \qquad H < \frac{1}{2},$$

$$\left(K^{*,-1}1_{[0,t]}(\cdot)\right)(s) = c(H)s^{\frac{1}{2}-H}\frac{d}{ds}\int_s^t u^{\frac{1}{2}-H}(u-s)^{-H+\frac{1}{2}}, \qquad H > \frac{1}{2}$$

To unify the notation, we write

$$\left(K^{*,-1}1_{[0,t]}(\cdot)\right)(s) = c(H)s^{\frac{1}{2}-H}\frac{d}{ds}\int_s^t u^{\frac{1}{2}-H}(u-s)^{-H+\frac{1}{2}}, \qquad H \in (0,1)$$

and we just observe that the constant $c(H)$ above is analytic with respect to H. Let us consider, for $v \leq t$ a function $A(v,t)$ such that

$$\int_v^t A(v,s)ds = \int_v^t \left(K^{*,-1}1_{[0,t]}(\cdot)\right)(s)K(s,v)ds.$$

Then, obviously,

$$Q_t = \int_0^t A(t,v)dZ_v.$$

On the other hand, it has been proved in [24] (see relations (3.4) and (3.5) therein) that for $H > \frac{1}{2}$,

$$Q_t^{KB} = \int_0^t A^{KB}(t,v)dZ_v^{KB}$$

with

$$A^{KB}(t,s) = c(H)(t^{2H-1} + s^{2H-1}).$$

Using the relations between Q and Q^{KB} and between Z and Z^{KB} (see Remark 4), it follows that, for every $H > \frac{1}{2}$, and $s < t$,

$$A(s,t) = c(H)\left[\left(\frac{s}{t}\right)^{\frac{1}{2}-H} + \left(\frac{t}{s}\right)^{\frac{1}{2}-H}\right]. \tag{59}$$

We show that the above relation (59) is true for $H < \frac{1}{2}$ as well. We use an argument inspired by [13], proof of Theorem 3.1. We observe that the functions

$$H \in (0,1) \to A(s,t) \text{ and } H \in (0,1) \to c(H)\left[\left(\frac{s}{t}\right)^{\frac{1}{2}-H} + \left(\frac{t}{s}\right)^{\frac{1}{2}-H}\right]$$

are analytic with respect to H and coincide on $(1/2, 1)$. Moreover, both are well-defined for every $H \in (0,1)$ (in fact it follows from [24] that A is well-defined for $H > \frac{1}{2}$ and it is more regular for $H \leq \frac{1}{2}$). To conclude (59) for every $H \in (0,1)$, we invoke the fact that if $f, g : (a,b) \to \mathbb{R}$ are two analytic functions and the set $\{x \in (a,b); f(x) = g(x)\}$ has an accumulation point in (a,b), then $f = g$.

As a consequence, (59) holds for every H and this shows that

$$\int_0^t Q_s dZ_s = \int_0^t Q_s^{KB} dZ_s^{KB} = c(H)\left(Z_t^{KB}\int_0^t r^{2H-1}dZ_r^{KB} - t\right)$$

and all the calculations contained in [24], Sections 3.2, 4 and 5 hold for every $H \in (0,1)$. ■

7.5 Proof of Theorem 4

7.5.1 Proof of Proposition 4

For conciseness, we only indicate how to establish one of the crucial estimates for this proposition, that the quantity

$$S_n := \frac{\sum_{m=0}^{n} \left(Q_m - \check{Q}_m\right)\left(Z_{m+1} - Z_m\right)}{\sum_{m=0}^{n} |Q_m|^2}$$

converges to 0 almost surely, and then only for $H < 1/2$. Since we want to show that S_n tends to 0 almost surely, and $\mathbf{P}$ and $\tilde{\mathbf{P}}$ share the same null sets, we may assume that Z is a Brownian motion, and X is a fractional Brownian motion adapted to Z's filtration.

Define the quantity

$$R_n = \sum_{j=0}^{n-1} (n-j)^{-H-1/2} j^{1/2-H} \int_j^{j+1} \left(b\left(X_j\right) - b\left(X_s\right)\right) ds.$$

This is related to S_n via the fact that $m^{H-1/2}|R_m| = Q_m - \check{Q}_m$. We claim that for any $\varepsilon > 0$, almost surely, for large m, that is $m \geq m_0$, $|R_m| \leq r_0 + m^{-H+1+\varepsilon} c_H \|b'\|$ where r_0 is a fixed random variable. This is sufficient to conclude that $\lim_n S_n = 0$. Indeed, we will see below (Section 7.5.3, Step 2, inequality (65)) that almost surely, for large n, $\sum_{m=0}^{n} |Q_m|^2 \geq n^2$. Then sum of all terms in the numerator of S_n for $m \leq m_0$, after having been divided by S_n's denominator, tend to 0 when $n \to \infty$. On the other hand, the IID terms $\{Z_{m+1} - Z_m\}_{m \in \mathbf{N}}$ are standard normal, so that one trivially proves that almost surely for $n \geq m_0$ (abusively using the same m_0 as above), up to some non-random universal constant c, $|Z_{m+1} - Z_m| \leq c\sqrt{\log m}$. It follows that the portion of S_n for $m \geq m_0$ is bounded above by $n^{-2} \sum_{m=0}^{n} m^{H-1/2} \left(r_0 + m^{-H+1+\varepsilon} c_H \|b'\|\right) \sqrt{\log m}$, which is itself bounded above by $(r_0 + c_H \|b'\|) n^{3/2+\varepsilon}$ which obviously tends to 0 as $n \to \infty$ as soon as $\varepsilon < 1/2$.

Now let us prove our claim on R_m. It is a known fact, which is obtained using standard tools from Gaussian analysis, or simply the Kolmogorov lemma, that for any $M \geq 1$

$$\mathbf{E}\left[\sup_{s,t\in[j,j+1]} |X_t - X_s|^M\right] \leq j^{HM}.$$

Indeed, this is easier to prove than Lemma 4. The usual application of the Borel-Cantelli lemma after Chebyshev's inequality for an M large enough, implies that for any $\alpha > H$, almost surely, for large j, $\sup_{s,t\in[j,j+1]} |X_t - X_s| \leq j^{\alpha}$. Consequently for any $\varepsilon > 0$,

$$\begin{aligned}
|R_m| &\leq 2\left\|b'\right\| \sum_{j=0}^{m_0} (n-j)^{-H-1/2} j^{1/2-H} \int_j^{j+1} |X_j - X_s|\, ds \\
&\quad + \left\|b'\right\| \sum_{j=m_0}^{n-1} (n-j)^{-H-1/2} j^{1/2-H} \sup_{s\in[j,j+1]} |X_j - X_s| \\
&= r_0 + n^{-2H} \left\|b'\right\| \sum_{j=m_0}^{n-1} (1-j/n)^{-H-1/2} (j/n)^{1/2-H} \sup_{s\in[j,j+1]} |X_j - X_s| \\
&\leq r_0 + n^{-2H} n^{H+\varepsilon} \sum_{j=m_0}^{n-1} (1-j/n)^{-H-1/2} (j/n)^{1/2-H} (j/n)^{H+\varepsilon} \\
&= r_0 + c_H n^{-H+\varepsilon+1} \left(1 + O\left(1/n\right)\right),
\end{aligned}$$

where the last estimate is in virtue of the Riemann sums for $\int_0^1 (1-x)^{-H-1/2} x^{1/2+\varepsilon} dx$.

7.5.2 Proof of Proposition 5

By our Theorems 2 and 3, it is of course sufficient to prove that

$$\lim_{n\to\infty} \left(\bar{\theta}_n - \theta_n\right) = 0.$$

In preparation for this, we first note that by classical properties for quadratic variations, and using our hypothesis, for large enough n, we have

$$\begin{aligned}\left|\langle B\rangle_n - \langle A\rangle_n\right| &= \left|\langle (B-A), (B+A)\rangle_n\right| \\ &\leq \left|\langle B+A\rangle_n\right|^{1/2} \left|\langle B-A\rangle_n\right|^{1/2} \\ &\leq \sqrt{2}n^{-\alpha} \left|\langle B\rangle_n\right|^{1/2} \left|\langle A\rangle_n + \langle B\rangle_n\right|^{1/2}. \end{aligned} \tag{60}$$

Now we prove that (60) implies almost surely,

$$\lim_{n\to\infty} \frac{\langle A\rangle_n}{\langle B\rangle_n} = 1. \tag{61}$$

Indeed let $x_n = \langle A\rangle_n / \langle B\rangle_n$. Then we can write

$$\begin{aligned}|x_n - 1| &= \frac{\left|\langle B\rangle_n - \langle A\rangle_n\right|}{\langle B\rangle_n} \\ &\leq \sqrt{2}n^{-\alpha} \left|\langle B\rangle_n\right|^{-1/2} \left|\langle A\rangle_n + \langle B\rangle_n\right|^{1/2} = c\sqrt{2}n^{-\alpha} \left|1 + x_n\right|^{1/2}.\end{aligned}$$

where c is a possibly random almost surely finite constant. Let $\varepsilon > 0$ be given; it is elementary to check that the inequality $(x-1)^2 \leq 2\varepsilon(x+1)$ is equivalent to $|x - (1+\varepsilon)| \leq \sqrt{4\varepsilon + \varepsilon^2}$. For us this implies immediately $|x_n - 1| \leq 6cn^{-\alpha}$, proving the claim (61).

Now we have

$$\theta_n - \bar{\theta}_n = \frac{A_n}{\langle A\rangle_n} - \frac{B_n}{\langle B\rangle_n} = \frac{A_n - B_n}{\langle B\rangle_n} + A_n \frac{\langle B\rangle_n - \langle A\rangle_n}{\langle A\rangle_n \langle B\rangle_n}. \tag{62}$$

Using (60) we have that the second term in (62) is bounded above in absolute value by

$$\sqrt{2}n^{-\alpha} \frac{A_n}{\langle A\rangle_n} \frac{\left|\langle A\rangle_n + \langle B\rangle_n\right|^{1/2}}{\left|\langle B\rangle_n\right|^{1/2}} = \sqrt{2}n^{-\alpha} \frac{A_n}{\langle A\rangle_n} \left(\frac{\langle A\rangle_n}{\langle B\rangle_n} + 1\right)^{1/2}.$$

By Theorems 2 and 3, $A_n / \langle A\rangle_n$ converges to the finite constant θ. By the limit (61), the last term in the above expression converges to 2, so that the entire expression converges to 0. Let k and γ be fixed positive values. For the first term in (62), using our hypotheses, by Chebyshev's and the Burkholder-Davis-Gundy inequalities, and from the expression of the semimartingales Z as $Z_t = \int_0^t Q_s dW_s + \theta \int_0^t Q_s ds$, we have

$$\begin{aligned}\mathbf{P}\left[|A_n - B_n|^k > n^{-k\gamma}\mathbf{E}\left[\langle B\rangle_n^k\right]\right] &\leq n^{\gamma k}\mathbf{E}^{-1}\left[\langle B\rangle_n^k\right]\mathbf{E}\left[|A_n - B_n|^k\right] \\ &\leq c(\theta)2^k n^{\gamma k} n^{-k\alpha}.\end{aligned}$$

Thus picking a positive value $\gamma < \alpha$ and choosing k large enough, by the Borell-Cantelli lemma, almost surely, for n large enough

$$|A_n - B_n| \leq n^{-\gamma}\mathbf{E}\left[\langle B\rangle_n^k\right]^{1/k} \leq \frac{1}{K} n^{-\gamma} \langle B\rangle_n ,$$

which finishes the proof of the proposition. ■

7.5.3 Proof of Theorem 4 (Steps 1 through 4)

In the entire proof below, $n_0(\omega)$ will denote a random, almost surely finite, integer; it may change from line to line, as it is introduced via various different applications of the Borel-Cantelli lemma, but one only needs to take the supremum of all such integers to have correct statements throughout.

Step 1. Bounding $|Q|^2$ below. Using only Condition (C'), we immediately get, for any $\gamma \in (0, 1/2 - H)$, for any large t,

$$\mathbf{P}\left[|Q_t| < t^{1/2-\gamma}\right] \leq K_b t^{-\gamma}.$$

To be able to apply the Borel-Cantelli lemma, we now let $t = n^A$ where n is an integer and A is a constant exceeding γ^{-1}. We then get, almost surely, for any $n > n_0(\omega)$,

$$|Q_{n^A}| > n^{A(1/2-\gamma)}. \tag{63}$$

We also bound other $Q_{m'}$'s that are in close proximity to Q_{n^A}. For any fixed integer $\bar{m}$, consider the set $I_{\bar{m}}$ of integers m' in the interval $[\bar{m}-\bar{m}^{4\gamma},\bar{m}]$, where γ is also assumed to be less than $1/4$. Then by Lemma 4 and Chebyshev's inequality, for any integer k,

$$\mathbf{P}\left[\sup_{m'\in I_{\bar{m}}}|Z_{\bar{m}}-Z_{m'}|>\bar{m}^{-\varepsilon}\right]\leq\frac{\bar{m}^{k\varepsilon}}{\bar{m}^{Hk(1-4\gamma)}}.$$

Thus for k large enough and $0<\varepsilon<H\left(1-4\gamma\right)$, by Borell-Cantelli's lemma, for $\bar{m}>m_0\left(\omega\right)$, for any integer $m'\in[\bar{m}-\bar{m}^{4\gamma},\bar{m}]$,

$$|Z_{m'}|>|Z_{\bar{m}}|-\bar{m}^{-\varepsilon},$$

from which we conclude, via the formula $Z_m=Q_m/\sqrt{m}$, that

$$|Q_{m'}|>|Q_{\bar{m}}|-\bar{m}^{1/2-\varepsilon}.$$

Certainly, if $\bar{m}$ is of the form n^A for large enough n, by choosing γ small enough, we obtain that the lower bound $\bar{m}^{1/2-\gamma}$ on $|Q_{\bar{m}}|^2$ obtained in (63) is dominant compared to $\bar{m}^{1/2-\varepsilon}$ for ε close to $H\left(1-4\gamma\right)$. Hence we get

$$|Q_{m'}|^2>|Q_{\bar{m}}|^2\left(1-\frac{\bar{m}^{1/2-\varepsilon}}{|Q_{\bar{m}}|}\right)^2\geq|Q_{\bar{m}}|^2/2. \tag{64}$$

Step 2. Bounding $\langle B\rangle$ from below. For n given, let n_1 be the largest integer such that $n_1^A\leq n<\left(n_1+1\right)^A$. Also assume n is large enough so that $n_1^A\geq n_0\left(\omega\right)$. Thus, applying (64) with $\bar{m}=n_1^A$,

$$\begin{aligned}\langle B\rangle_m &\geq\sum_{m=0}^{n_1^A}|Q_m|^2\geq\left|Q_{n_1^A}\right|^2+\sum_{m'=n_1^A-\left(n_1^A\right)^{4\gamma}}|Q_{m'}|^2\\ &\geq\left|Q_{n_1^A}\right|^2+2^{-1}\sum_{m'=n_1^A-\left(n_1^A\right)^{4\gamma}}\left|Q_{n_1^A}\right|^2\geq\left|Q_{n_1^A}\right|^2\left(n_1^A\right)^{4\gamma}.\end{aligned}$$

We can now invoke (63) to say that almost surely, for $n>n_0\left(\omega\right)^A$

$$\langle B\rangle_n\geq\left(n_1\right)^{A(2-2\gamma)}\left(n_1^A\right)^{4\gamma}=2^{-1}\left(n_1\right)^{A(2+2\gamma)}.$$

Given that we may write $n_1^A\left(1+n_1^{-1}\right) > n$, so that $n_1^A > n/2$, we can finally conclude that

$$\langle B\rangle_n \geq \frac{1}{2^{1+2A(1+\gamma)}} n^{2+2\gamma}. \tag{65}$$

Step 3. Bounding $\langle A-B\rangle$'s terms from above. We may generically bound the general term of $\langle A-B\rangle_n$:

$$\begin{aligned}\int_m^{m+1} |Q_s - Q_m|^2\, ds &= \int_m^{m+1} \left|\sqrt{s}Z_s - \sqrt{m}Z_m\right|^2 ds \\ &\leq 2\,(m+1)\int_m^{m+1} |Z_s - Z_m|^2\, ds + 2\,|Z_m|^2 \int_m^{m+1} \left(\sqrt{s}-\sqrt{m}\right)^2 ds \\ &\leq 2\,(m+1)\int_m^{m+1} |Z_s - Z_m|^2\, ds + 2\,|Z_m|^2/m. \end{aligned} \tag{66}$$

We begin by dealing with the first term in (66): by Lemma 4 for any $M>0$,

$$\mathbf{P}\left[\sup_{s\in[m,m+1]} |Z_m - Z_s|^2 > m^{-2\delta}\right] \leq \frac{1}{m^{M(H-\delta)}}.$$

Hence for $\delta < H$, for M large enough, by the Borel-Cantelli lemma, for $m > m_0(\omega)$,

$$\int_m^{m+1} |Z_m - Z_s|^2\, ds \leq \sup_{s\in[m,m+1]} |Z_m - Z_s|^2 \leq m^{-2\delta}. \tag{67}$$

For the second term in (66), which involves Z_m, we note that the hypothesis b' bounded implies that for some constant c_b, $|b(x)| \leq c_b(1+|x|)$. Thus by (26),

$$|Z_m| = \left|\int_0^1 \mu(dr)\, \frac{b(\omega_{rm})}{m^H}\right| \leq c_{b,H} + c_b m^{-H}\int_0^1 |\omega_{mr}|\, dr.$$

The random variable $Y_m = m^{-H}\int_0^1 |\omega_{mr}|\, dr$ is equal in distribution to a sub-Gaussian random variable: let y_H be its mean, which only depends on H. Therefore, there exists a number σ_H which also only depends on H such that for each m and $x>0$,

$$\mathbf{P}\left[|Y_m - y_H| > x\right] \leq 2\exp\left(-\frac{(x-y_H)}{2\sigma_H^2}\right).$$

For $x = m^{\alpha}$, this translates as the existence of a constant $y_{b,H}$ depending only on H and b such that

$$\mathbf{P}\left[|Z_m| > \mu_{b,H} + \sqrt{4c_b^2\sigma_H^2 \log m}\right] \leq 2\exp\left(-\frac{4\sigma_H^2 \log m}{2\sigma_H^2}\right) = \frac{1}{m^2},$$

so that there exists a constant $c_{b,H}$ depending only on b and H such that by the Borel-Cantelli lemma, almost surely, for $m > m_0(\omega)$,

$$|Z_m|^2 \leq c_{b,H} \log m. \tag{68}$$

Plugging (67) and (68) into (66), we conclude that for any $\delta < H$, almost surely, for $m > m_0(\omega)$,

$$\int_m^{m+1} |Q_s - Q_m|^2\, ds \leq c_{b,H}\frac{\log m}{m} + m^{1-2\delta} \leq 2m^{1-2\delta}.$$

Step 4. Conclusion. From the formula $\langle A - B\rangle_n = \sum_{m=0}^{n-1}\int_m^{m+1} |Q_s - Q_m|^2\, ds$, using the last estimate of the previous step, we get

$$\langle A - B\rangle_n \leq \sum_{m=0}^{m_0(\omega)} \int_m^{m+1} |Q_s - Q_m|^2\, ds + 2n^{2-2\delta}.$$

From the final estimate (65) of Step 2, we may now write almost surely

$$\frac{\langle A - B\rangle_n}{\langle B\rangle_n} \leq \frac{\sum_{m=0}^{m_0(\omega)}\int_m^{m+1} |Q_s - Q_m|^2\, ds}{n^{2+2\gamma}} + \frac{1}{n^{2(\gamma+\delta)}}.$$

Hence the first statement of Proposition 5 is established for any $\alpha < 2H$. ■

References

[1] Y. Aït Sahalia (2002). Maximum likelihood estimation of discretely sampled diffusions: a closed form approximation approach. *Econometrica*, **70**, pag. 223-262.

[2] E. Alos, O. Mazet and D. Nualart (2001). Stochastic calculus with respect to Gaussian processes. *Annals of Probability*, **29**, 766-801.

[3] J.-M. Bardet; G. Lang; G. Oppenheim; A. Philippe; S. Stoev. ; M. Taqqu (2003). Semi-parametric estimation of the long-range dependence parameter : a survey. In *Theory and applications of long-range dependence*, 557–577, Birkhäuser.

[4] A. Beskos, O. Papaspiliopoulos, G. Roberts and P. Fearnhead (2006). Exact and computationally efficient likelihood-based inference for discretely observed diffusion processes. *J.R. Statistical Soc. B* **68** (2), 1-29.

[5] B. Bibby and M. Sorensen (1995). Martingale estimation functions for discretely observed diffusion processes. *Bernoulli*, **1**, 17-39.

[6] J.P. Bishwal; A. Bose (2001). Rates of convergence of approximate maximum likelihood estimators in the Ornstein-Uhlenbeck process. *Comput. Math. Appl.* **42**, no. 1-2, 23–38.

[7] B. Boufoussi and Y. Ouknine (2003). On a stochastic equation driven by a fBm with discontinuous drift. *Electronic Comm. in Probab.*, **8**, 122-134.

[8] I.V. Basawa, B.L.S. Prakasa Rao (1980). *Statistical Inference for Stochastic Processes.* Academic press, London.

[9] P. Cheridito, H. Kawaguchi, M. Maejima (2003). Fractional Ornstein-Uhlenbeck processes. *Electronic Journal of Probability*, **8**, paper no. 3, 1-14.

[10] E.M. Cleur (2001). Maximum likelihood estimates of a class of one-dimensional stochastic differential equation models from discrete data. *J. Time Ser. Anal.* **22**, no. 5, 505–515.

[11] J.-F. Coeurjolly (2005). *L-type estimators of the fractal dimension of locally self-similar Gaussian processes.* Preprint, online at http://hal.ccsd.cnrs.fr/docs/00/03/13/77/PDF/robustHurstHAL.pdf.

[12] D. Dacunha-Castelle and D. Florens-Zmirou (1986). Estimation of the coefficients of a diffusion from discrete observations. *Stochastics* **19**, 263-284.

[13] L. Decreusefond and A.S. Ustunel (1999). Stochastic analysis of the fractional Brownian motion. *Potential Analysis*, **10**, 177-214.

[14] B. Djehiche and M. Eddahbi (2001). Hedging options in market models modulated by fractional Brownian motion. *Stochastic Analysis and Applications*, **19** (5), 753-770.

[15] O. Elerian, S. Chib, N. Shephard (2001). Likelihood inference for discretely observed nonlinear diffusions. *Econometrica* **69**, 959-993.

[16] X. Fernique (1974): *Régularité des trajectoires de fonctions aléatoires gaussiennes.* Lecture Notes in Mathematics (St. Flour), 480, pp. 2-95.

[17] A. Friedman (1975). *Stochastic differential equations and applications.* Academic Press.

[18] C. Gourieroux, A. Monfort and E. Renault (1993). Indirect Inference. *Econometrics* **8**, 85-118.

[19] Y. Hu and B. Oksendhal (2003). Fractional white noise calculus and applications to finance. *IDAQP*, **6** (1), 1-32.

[20] Yu. A. Kutoyants (1977). On a property of estimator of parameter of trend coefficient. *Izv. Akad. Nauk Arm. SSR.*, *Mathematika*, **12**, 245-251.

[21] Yu. A. Kutoyants (1977). Estimation of the trend parameter of a diffusion process. *Theory Probab. Appl.*, **22**, 399-405.

[22] Yu. A. Kutoyants (1980). *Parameter Estimation for Stochastic Processes*, Heldermann, Berlin, 1984. (Russian edition 1980).

[23] Yu. A. Kutoyants (2004). *Statistical Inference for Ergodic Diffusion Processes.* Springer Series in Statistics. Springer.

[24] M.L Kleptsyna and A. Le Breton (2002). Statistical analysis of the fractional Ornstein-Uhlenbeck type processes. *Statistical inference for stochastic processes*, **5**, 229-248.

[25] M. Kleptsyna, A. Le Breton and M.C. Roubaud (2000). Parameter estimation and optimal filtering for fractional type stochastic systems. *Statistical inference for stochastic processes*, **3**, 173-182.

[26] A.Kukush, Y. Mishura and E. Valkeila (2005). Statistical inference with fractional Brownian motion. *Statistical inference for stochastic processes*, **8**, 71-93.

[27] A. Le Breton (1976). On continuous and discrete sampling for parameter estimation in diffusion type processes. *Mathematical Programming studies*, **5**, 124-144.

[28] R.S. Liptser and A.N. Shiryayev (1978). *Statistics of random processes II. Applications* Springer-Verlag.

[29] P. Malliavin (1976): *Stochastic calculus of variations and hypoelliptic operators.* Proc. Inter. Symp. on Stoc. Diff. Eqs., 195-273. Wiley.

[30] I. Norros, E. Valkeila and J. Virtamo (1999). An elementary approach to a Girsanov formula and other analytical results on fractional Brownian motion. *Bernoulli*, **5** (4), 571-587.

[31] D. Nualart (1995). *The Malliavin calculus and related topics.* Springer-Verlag.

[32] D. Nualart and Y. Ouknine (2002). Regularization of differential equations by fractional noise. *Stoc. Proc. Appl.*,**102**, 103-116.

[33] A.R. Pedersen, (1995). A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations. *Scand. J. Statist.* **22**, no. 1, 55–71.

[34] R. Poulsen (1999). *Approximate maximum likelihood estimation of discretely observed diffusion processes.* Center for Analytical Finance, working paper 29.

[35] D. Revuz and M. Yor (1999). *Continuous Martingales and Brownian Motion. Third edition.* Springer-Verlag.

[36] B.L.S. Prakasa Rao (2003). Parameter estimation for linear stochastic differential equations driven by fractional Brownian motion. *Random Oper. Stoc. Eqs.*, **11** (3), 229-242.

[37] B.L.S. Prakasa Rao (1999). *Statistical Inference for Diffusion type processes.* Oxford University Press.

[38] H. Sorensen (2004). *Parametric inference for diffusion processes observed at discreted points time: a survey.* Int. Stat. Rev. **72**, 337-354.